24th Annual National Conference on Managing Environmental Quality Systems

8:30 – 12:00 TUESDAY, APRIL 12TH - A.M. Stockholder Meetings

12:00 – 4:30 TUESDAY, APRIL 12TH
Opening Plenary (Salons A-H)

- Opening Address
  - Reggie Cheatham, Director, OEI Quality Staff, EPA
  - Linda Travers, Principal Deputy Assistant Administrator, OEI, EPA
- Invited Speakers
  - Tom Huetteman, Deputy Assistant Regional Administrator, EPA Region 9
  - John Robertus, Executive Officer of San Diego Regional Water Quality Control Board, Region 9
- Keynote Address
  - Thomas Redman, President, Navesink Consulting Group
- Panel Sessions
- **Value of the Data Quality Act—Perspectives from OMB, Industry, and EPA (VDQA)**
  - Nancy Beck, OMB
  - Jamie Conrad, American Chemistry Council
  - Reggie Cheatham, Director, OEI Quality Staff, EPA
- **Wadeable Streams: Assessing the Quality of the Nation's Streams (WS)**
  - Margo Hunt, Panel Moderator
  - Mike Shapiro, Deputy Assistant Administrator, Office of Water
  - Steve Paulsen, Research Biologist, ORD

8:30 – 10:00 WEDNESDAY, APRIL 13TH
Environmental Measures (EM) (Salons A-C) *Chair: L. Bradley, EPA*

- Data Error Reduction by Automation throughout the Data Workflow Process (A. Gray, EarthSoft, Inc.)
- Analytical Approaches to Meeting New Notification Levels for Organic Contaminants in Calif. (D.Wijekoon, Calif. DHS)
- Streamlining Data Management and Communications for the Former Walker AFB Project (R. Amano, Lab Data Consultants, Inc.)

Quality System Implementation in the Great Lakes Program (QSI-GLP) (Salon D) *Chair: M. Cusanelli, EPA*

- GLNPO's Quality System Implementation for the New "Great Lakes Legacy Act for Sediment Remediation"(L. Blume, EPA)
- Black Lagoon Quality Plan Approval by GLNPO, MDEQ, ERRS, and USACE (J. Doan, Environmental Quality Management, Inc.)
- Remediation of the Black Lagoon Trenton Channel . . . Postdredging Sampling & Residuals Analysis (J. Schofield, CSC)

Quality Systems Models (QSM) (Salons F-H) *Chair: G. Johnson, EPA*

- Improving E4 Quality System Effectiveness by Using ISO 9001: 2000 Process Controls (C. Hedin, Shaw Environmental)

Applications of Novel Techniques to Environmental Problems (ANTEP) (Salon E) *Chair: B. Nussbaum, EPA*

- On Some Applications of Ranked Set Sampling (B. Sinha, University of Maryland)
- Combining Data from Many Sources to Establish Chromium Emission Standards (N. Neerchal, University of Maryland)
- Estimating Error Rates in EPA Databases for Auditing Purposes (H. Lacayo, Jr., EPA)
- Spatial Population Partitioning Using Voronoi Diagrams For Environmental Data Analysis (A. Singh, UNLV)

Ambient Air Session I (Sierra 5&6) *Chair: M.Papp, EPA*

- Changes and Improvements in the Ambient Air Quality Monitoring Program Quality System (M. Papp, EPA)
- Guidance for a New Era of Ambient Air Monitoring (A. Kelley, Hamilton County DES)
- Environmental Monitoring QA in Indian Country (M. Ronca-Battista, Northern Arizona University)
- Scalable QAPP IT Solution for Air Monitoring Programs (C. Drouin, Lake Environmental Software)

10:30 – 12:00 WEDNESDAY, APRIL 13TH

Environmental Laboratory Quality Systems (ELQS) (Salons A-C) *Chair: L. Bradley, EPA*

- A Harmonized National Accreditation Standard: The Next Step for INELA Field Activities (D. Thomas, Professional Service Industries, Inc.)
- Development of a Comprehensive Quality Standard for Environmental Laboratory Accreditation (J. Parr, INELA)
- Advanced Tracking of Laboratory PT Performance and Certification Status with Integrated Electronic NELAC-Style Auditing Software (T. Fitzpatrick, Lab Data Consultants, Inc.)

Performance Metrics (PM) (Salon D) *Chair: L. Doucet, EPA*

- Formulating Quality Management Metrics for a State Program in an Environmental Performance Partnership Agreement (P. Mundy, EPA)
- How Good Is "How Good Is?" (Measuring QA) (M. Kantz, EPA)
- Performance-Based Management (J. Santillan, US Air Force)

Quality Assurance Plan Guidance Initiatives (QAPGI) (Salons F-H) *Chair: A. Batterman, EPA*

- A CD-ROM Based QAPP Preparation Tool for Tribes (D. Taylor, EPA)
- Military Munitions Response Program Quality Plans (J. Sikes, U.S. Army)

Ask a Statistician: Panel Discussion (Salon E) *Moderator: B. Nussbaum, EPA* Panelists:

- Mike Flynn, Director, Office of Information Analysis and Access, OEI, EPA
- Reggie Cheatham, Director, Quality Staff, OEI, EPA
- Tom Curran, Chief Information Officer, OAQPS, EPA
- Diane Harris, Quality Office, Region 7, EPA
- Bill Hunt, Visiting Senior Scientist, North Carolina State University (NCSU)
- Rick Linthurst, OIG, EPA

Ambient Air Session II (Sierra 5&6) *Chair: M. Papp, EPA*

- National Air Toxics QA System and Results of the QA Assessment (D. Mikel, EPA)
- Technical System Audits (TSAs) and Instrument Performance Audits (IPAs) of the National Air Toxics Trends Stations (NATTS) and Supporting Laboratories (S. Stetzer Biddle, Battelle)
- Interlaboratory Comparison of Ambient Air Samples (C. Pearson, CARB)
- Developing Criteria for Equivalency Status for Continuous PM2.5 Samplers (B. Coutant, Battelle)

1:00 – 2:30 WEDNESDAY, APRIL 13TH

Environmental Laboratory Quality (ELQ) (Salons A-C) *Chair: L. Doucet, EPA*

- Environmental Laboratory Quality Systems: Data Integrity Model and Systematic Procedures (R. DiRienzo, DataChem Laboratories, Inc.)
- The Interrelationship of Proficiency Testing, Interlaboratory Statistics and Lab QA Programs (T. Coyner, Analytical Products Group, Inc.)
- EPA FIFRA Laboratory Challenges and Solutions to Building a Quality System in Compliance with International Laboratory Quality Standard ISO 17025 (A. Ferdig, Mich. Dept. of Agriculture)

Performance—Quality Systems Implementation (P-QSI) (Salon D) *Chair: A. Belle, EPA*

- Implementing and Assessing Quality Systems for State, Tribal, and Local Agencies (K. Bolger, D. Johnson, L. Blume, EPA)

1:00 – 2:30 WEDNESDAY, APRIL 13TH (continued)

Quality Initiatives in the EPA Office of Environmental Information (QI-OEI) (Salons F-H) *Chair: J. Worthington, EPA*

- Next Generation Data Quality Automation in EPA Data Marts (P. Magrogan, Lockheed)
- The Design and Implementation of a Quality System for IT Products and Services (J. Scalera, EPA)
- Data Quality is in the Eyes of the Users: EPA's Locational Data Improvement Efforts (P. Garvey, EPA)

A Win-Win-Win Partnership for Solving Environmental Problems (W3PSEP) (Salon E) *Co-Chairs: W. Hunt, Jr. and K. Weems, NCSU*

- Overview of Environmental Statistics Courses at NCSU (B. Hunt, NCSU Statistics Dept.)
- Overview of the Environmental Statistics Program at Spelman College (N. Shah, Spelman)
- Student presentations: H. Ferguson and C. Smith of Spelman College; C. Pitts, B. Stines and J. White of NCSU

Ambient Air Session III (Sierra 5&6) *Chair: M. Papp, EPA*

- Trace Gas Monitoring for Support of the National Air Monitoring Strategy (D. Mikel, EPA)
- Comparison of the Proposed Versus Current Approach to Estimate Precision and Bias for Gaseous Automated Methods for the Ambient Air Monitoring Program (L. Camalier, EPA)
- Introduction to the IMPROVE Program's New Interactive Web-based Data Validation Tools (L. DeBell, Colorado State University)
- The Role of QA in Determination of Effects of Shipping Procedures for PM2.5 Speciation Filters (D. Crumpler, EPA)

3:00 – 4:30 WEDNESDAY, APRIL 13TH

Topics in Environmental Data Operations (TEDO) (Salons A-C) *Chair: M. Kantz, EPA*

- Ethics in Environmental Operations: It's More Than Just Lab Data (A. Rosecrance, Laboratory Data Consultants, Inc.)
- QA/QC of a Project Involving Cooperative Agreements, IAGs, Agency Staff and Contracts to Conduct the Research (A. Batterman, EPA)
- Dealing with Fishy Data: A Look at Quality Management for the Great Lakes Fish Monitoring Program (E. Murphy, EPA)

Quality System Development (QSD) (Salon D) *Chair: A. Belle, EPA*

- Development of a QA Program for the State of California (B. van Buuren, Van Buuren Consulting, LLC)
- Integrating EPA Quality System Requirements with Program Office Needs for a Practical Approach to Assuring Adequate Data Quality to Support Decision Making (K. Boynton, EPA)
- Introducing Quality System Changes in Large Established Organizations (H. Ferguson, EPA)

Auditor Competence (AC) (Salons F-H) *Chair: K. Orr, EPA*

- Determining the Competence of Auditors (G. Johnson, EPA)

To Detect or Not Detect—What Is the Problem? (TDND) (Salon E) *Chair: J. Warren, EPA*

- A Bayesian Approach to Measurement Detection Limits (B. Venner)
- The Problem of Statistical Analysis with Nondetects Present (D. Helsel, USGS)
- Handling Nondetects Using Survival Anal.(D. Helsel, USGS)
- Assessing the Risk associated with Mercury: Using ReVA's Webtool to Compare Data, Assumptions and Models (E. Smith, EPA)

Ambient Air Session IV (Sierra 5&6) *Chair: M. Papp, EPA*

- Status and Changes in EPA Infrastructure for Bias Traceability to NIST (M. Shanis, EPA)
- Using the TTP Laboratory at Sites with Higher Sample Flow Demands (A. Teitz, EPA)

5:00 – 6:00 PM WEDNESDAY, APRIL 13TH

EPA SAS Users Group Meeting Contact: Ann Pitchford, EPA

8:30 – 10:00 THURSDAY, APRIL 14TH

Evaluating Environmental Data Quality (EEDQ) *(Salons A-C) Chair: M. Kantz, EPA*

- QA Documentation to Support the Collection of Secondary Data (J. O'Donnell, Tetra Tech, Inc.)
- Staged Electronic Data Deliverable: Overview and Status (A. Mudambi, EPA)
- Automated Metadata Reports for Geo-Spatial Analyses (R. Booher, INDUS Corporation)

Satellite Imagery QA (SI-QA) (Salon D) *Chair: M. Cusanelli, EPA*

- Satellite Imagery QA Concerns (G. Brilis and R. Lunetta, EPA)

Information Quality Perspectives (IQP) (Salons F-H) *Chair: J. Worthington, EPA*

- A Body of Knowledge for Information and Data Quality (J. Worthington, L. Romero Cedeno, EPA)
- Information as an Environmental Technology – Approaching Quality from a Different Angle (K. Hull, Neptune and Co.)

To Detect or Not Detect—What Is the Answer? (TDND) (Salon E) *Chair: A. Pitchford, EPA, Co-Chair: W. Puckett, EPA*

- Using Small Area Analysis Statistics to Estimate Asthma Prevalence in Census Tracts from the National Health Interview Survey (T. Brody, EPA)
- Logistical Regression and QLIM Using SAS Software (J. Bander, SAS)
- Bayesian Estimation of the Mean in the Presence of Nondetects (A. Khago, University of Nevada)

Ambient Air Workgroup Meeting (Sierra 5&6) *Contact: Mike Papp, EPA*
NOTE: This is an all-day, closed meeting.

10:30 – 12:00 THURSDAY, APRIL 14TH

Environmental Data Quality (EDQ) (Salons A-C) *Chair: V. Holloman, EPA*

- Assessing Environmental Data Using External Calibration Procedures (Y. Yang, CSC)
- Groundwater Well Design Affects Data Representativeness: A Case Study on Organotins (E. Popek, Weston Solutions)

Information Quality and Policy Frameworks (IQPF) (Salons F-H) *Chair: L. Doucet, EPA*

- Modeling Quality Management System Practices to an Organization's Performance Measures (J. Worthington, L. Romero Cedeño, EPA)
- Development of a QAPP for Agency's Portal (K. Orr, EPA)
- Discussion of Drivers and Emerging Issues, Including IT, That May Result in Revisions to EPA's Quality Order and Manual (R. Shafer, EPA)

Office of Water; Current Initiatives (OW) (Salon D) *Chair: D. Sims, EPA*

- Whole Effluent Toxicity--The Role of QA in Litigation (M. Kelly, EPA, H. McCarty, CSC)
- Review of Data from Method Validation Studies: Ensuring Results Are Useful Without Putting the Cart Before the Horse (W. Telliard, EPA, H. McCarty, CSC)
- Detection and Quantitation Concepts: Where Are We Now? (Telliard, Kelly, and McCarty)

Sampling Inside, Outside, and Under (SIOU) (Salon E) *Chair: J. Warren, EPA*

- VSP Software: Designs and Data Analyses for Sampling – Contaminated Buildings (B. Pulsipher, J. Wilson, Pacific Northwest National Laboratory , R. O. Gilbert)
- Incorporating Statistical Analysis for Site Assessment into a Geographic Information System (D. Reichhardt, MSE Technology Applications, Inc.)
- The OPP's Pesticide Data Program Environmental Indicator Project (P. Villanueva, EPA)

1:00 – 2:30 THURSDAY, APRIL 14TH

Information Management (Salons A-C) *Chair: C. Thoma, EPA*

- Achieve Information Management Objectives by Building and Implementing a Data Quality Strategy (F. Dravis, Firstlogic)

UFP Implementation (Salon D) *Chair: D. Sims, EPA*

- Implementing the Products of the Intergovernmental DQ Task Force: The UFP QAPP (R. Runyon, M. Carter, EPA)
- Measuring Performance: The UFP QAPP Manual (M. Carter, EPA, C. Rastatter, VERSAR)

Quality Systems Guidance and Training Developments (QSG) (Salons F-H) *Chair: M. Kantz, EPA*

- A Sampling and Analysis Plan Guidance for Wetlands Projects (D. Taylor, EPA)
- My Top Ten List of Important Things I Do as an EPA QA and Records Manager (T. Hughes, EPA)
- I'm Here---I'm Free----Use Me! Use Me!—Secondary Use of Data in Your Quality System (M. Kantz, EPA)

Innovative Environmental Analyses (IEA) (Salon E) *Chair: M. Conomos, EPA*

- Evaluation of Replication Methods between NHANES 1999-2000 and NHANES 2001-2002 (H. Allender, EPA)
- Assessment of the Relative Importance of the CrEAM Model's Metrics (A. Lubin, L. Lehrman, and M. White, EPA)
- Statistical Evaluation Plans for Compliance Monitoring Programs (R. Ellgas, Shaw Environmental, Inc.; J. Shaw, EMCON/OWT, Inc.)

A Harmonized National Accreditation Standard: The Next Step for INELA Field Activities

By Dawn D. Thomas, ASQ CQM

The original charter of the National Environmental Laboratory Accreditation Conference (NELAC), when established in the early 1990's, was to "foster the generation of environmental laboratory data of known and documented quality through the development of national performance standards for environmental laboratories". However, it has been generally recognized within the environmental community, over the years, that the quality of environmental laboratory data can only be assured if minimum performance standards exist for field sampling and measurement activities – the "front-end" of the environmental data generation process. To assure the production of environmental data that are scientifically valid and can be used with a high degree of confidence by the end-user, control of environmental laboratory analytical processes and field sampling and measurement processes are of equal and significant importance. Accordingly, in July 1998, the Constitution of NELAC was amended to reflect the growing interest of many stakeholders to expand its scope to include both field sampling and measurement activities. Subsequent to this Constitutional amendment, the Field Activities Committee was officially established in 1999 as a NELAC standing committee responsible for the development of performance standards applicable to those organizations performing field sampling and measurement activities.

In July 2002, Chapter 7, *Field Activities Standard*, was added to the NELAC Standard to address minimum quality and technical requirements for field sampling and measurement activities. The initial draft of this chapter excerpted selected verbiage from Chapter 5, *Quality Systems*, of the NELAC laboratory standard and did not specifically address other accreditation components (e.g., proficiency testing (PT), on-site assessment, and accreditation process) or requirements for sampling specific environmental matrices. In 2003, NELAC divested itself of the environmental standards development process and the Institute for National Environmental Laboratory Accreditation (INELA), a consensus based standards development organization, was formed. Within this organization, the INELA Field Activities Committee (FAC) was established to continue the standards development work for an accreditation program designed specifically for field sampling and measurement organizations (FSMO).

Objective and Goals

The primary objective of the INELA FAC is "to develop and maintain consensus accreditation standards and guidance materials for organizations engaged in environmentally related field sampling and measurement activities, consistent with regulatory and industry-specific requirements". Its long-range focus is to replace the 2002 NELAC *Field Activities Standard* (Chapter 7) with an INELA stand-alone, FSMO-specific accreditation standard(s) that meets the following goals:

- Encompasses broad scope and wide ranging applicability;
- Based on internationally recognized standards for competency (ISO/IEC 17025) and conformity assessment (ISO/IEC 17011);
- NOT prescriptive in nature, allowing for the development of FSMO-specific policies and procedures; and

- Effectively supported by sound guidance.

Broad Scope and Applicability

If the INELA FAC is to meet its objective of establishing performance standards for those collecting samples and conducting on-site measurements for improved environmental data quality, then the standard must be wide-ranging in scope and applicability to support existing and future state/federal environmental regulations governing field sampling and measurement activities. To this end, a primary goal of the INELA Field Activities Committee is to develop an accreditation standard (or series of standards) that will apply to organizations performing field activities for a wide variety of sampling and measurement media such as air, biological, water, soil, waste, and radiological. Due to the nuances, specific to each media, a "one size fits all" approach to standards development is not appropriate. Accordingly, the FAC has engaged field sampling and measurement "media experts" to collaborate on the development of customized, media-specific FSMO accreditation standards. The development of custom field standards for water and air are the current focus of the committee.

ISO Foundation

It is the consensus viewpoint of the Field Activities Committee that the common denominator, or foundation, for the custom, media-specific INELA FSMO accreditation standard(s) must be ISO/IEC 17025, *General Requirements for the Competence of Testing and Calibration Laboratories* and ISO/IEC 17011 (soon to replace ISO/IEC Guide 58), *Conformity Assessment – General Requirements for Accreditation Bodies Accrediting Conformity Assessment Bodies*. Using this approach to standards development, the role of the INELA FAC will be to utilize its "media experts" to determine **how** to best apply these generic International Standards for a particular area of accreditation (e.g., field activities – water). The INELA FAC "application" of these International Standards, for each sampling and measurement media, will include, but will not be limited to, provisions for additional requirements, exclusion of specified ISO requirements due to applicability concerns, and clarifications and interpretations of various ISO requirements. Using ISO as the foundation for custom-built FSMO accreditation standards facilitates harmonization of individual field standards specific to each sampling and measurement media.

Non-Prescriptive Standards Development

Although sampling has, historically, been recognized as a major contributor to the overall measurement error, many organizations performing field sampling and measurement activities today are not currently subject to rigorous and prescriptive quality system requirements, accreditation, or routine oversight. Accordingly, the committee consensus was to take a practical and realistic first step towards improved environmental data quality by establishing an accreditation standard, based on internationally recognized standards, which are minimally prescriptive to provide a high degree of flexibility for the FSMO when implementing the standard requirements. Simply stated, applying this "less is better" approach, the FSMO will be able to craft policies and procedures, which meet the intent of the INELA standard, but are practical, functional and, most importantly, implement-able. The INELA FAC believes that if

the resulting field accreditation standards cannot be effectively implemented by all parties affected, large and small, public and private, due to overly prescriptive requirements, then we, as a committee, have not successfully completed our mission for improving data quality for better decisions.

Sound Guidance

To support the "less is better" approach to standards development and to facilitate successful implementation by all FSMO impacted by the standard, the development of appropriate implementation guidance tools is a key component for realizing an improved outcome – sound and defensible data quality for better decisions. This is the long-term focus of the INELA Field Activities Committee - to "show the way" by providing the necessary guidance and support for standards implementation. Several of the many benefits associated with this INELA service to the environmental community include:

- Acceleration of the FSMO "learning curve" associated with "something new", keeping in mind that many FSMO have not been subject to quality system/accreditation program requirements, historically;
- Improved "buy-in" by minimizing the costs associated with implementation of a new and comprehensive accreditation standard; and
- Consistency of standards interpretation and implementation.

Accomplishments

These goals for standards development, as discussed in the previous sections, have evolved over a period of two (2) years as a result of the diligent work and "outside the box" thinking of the INELA FAC. The accomplishments, which follow in this section, have contributed greatly to the refocusing of the laboratory community (regulators and those regulated) on the importance of field sampling and measurement and its role, as the "front-end" portion of the environmental data generation process.

To facilitate the development of media-specific field standards, the committee has been very active in outreach activities to engage more stakeholders – the "media experts" - in the standards development process. The INELA FAC has grown from less then ten (10) members in 2003 to more than thirty (30) participating members today. The committee has also worked to achieve balance of membership, necessary for a consensus standards development organization, with representation from government and municipal agencies; engineering and environmental consulting firms, analytical laboratories and industry. Participation in national/regional conferences and collaboration with other organizations representing specific stakeholder groups will continue to be a focus for the INELA FAC. The committee's success in developing sound field accreditation standards depends on the continuation of these outreach activities.

Consistent with committee direction to develop "applications" of the ISO/IEC 17025 and 17011 standards, a generic (not specific to any one media) application of the ISO/IEC 17025 standard has been completed and will be utilized by the "media experts" to guide the development of media-specific field accreditation standards. This generic application of ISO/IEC 17025 was

affirmed by the INELA membership in late 2004. Additionally, the groundwork, in the form of a consensus-based conceptual model, for the application of the ISO/IEC 17011 standard was completed and presented at the INELA Accreditation Forum in Charleston, South Carolina last summer. Building on these endeavors, workgroups have been established and are tasked with producing the first Working Draft Standards for a generic application of 17011 and a media-specific (water) application of 17025 by the summer of 2005.

A great deal has been accomplished but there is more work to do.

Next Steps

To achieve its on-going objective "to develop and maintain consensus accreditation standards and guidance materials for organizations engaged in environmentally related field sampling and measurement activities, consistent with regulatory and industry specific requirements", the INELA Field Activities Committee must effectively meet certain challenges. They are:

- To know, engage and understand the needs of all stakeholders who will be, ultimately, impacted by the standard(s).
- To know, engage and understand the needs of all potential clients, those who will adopt and implement such a standard(s).
- Finding a consensus viewpoint to the question of *what makes for good quality* to achieve consistent application of the ISO/IEC 17025 and 17011 standards for harmonized individual media-specific field accreditation standards.

With its new approach to standards development, the INELA FAC also has an opportunity to help chart the future path of INELA, as a standards development organization. At the 2004 INELA Summer Forum in Charleston, South Carolina, the INELA Board of Directors expressed their desire for INELA membership to seriously consider a restructuring of the NELAC laboratory standard to better meet the needs of stakeholders, existing and potential clients, and to achieve the desire growth into other areas of accreditation. There are a number of proposals for this restructuring initiative currently being considered by the INELA Board.

One of the proposals being considered has been developed by the INELA FAC, which details an approach to standard restructuring, consistent with the approach being taken for the development of media-specific field accreditation standards. This proposal has been designed to:

- Align with the INELA Strategic Plan.
- Provide a flexible framework for the development of harmonized accreditation standards in new areas such as Homeland Security.
- Positively impact a wide range of stakeholders.
- Appeal to accrediting authorities, regulators, private sector groups interested in adopting and implementing uniform standards of accreditation.
- Assure the production of scientifically valid data that can be used with a high degree of confidence by the end user.

The INELA Field Activities Committee is committed to the development of field accreditation standards using the approach detailed in this paper and strongly believes that this approach can be effectively used for the development of new INELA standards in other areas of accreditation as well. To meet the current challenges and to adequately address the complexities of the field sampling and measurement "world", the committee must continue to focus its energies on thinking "outside the box", encouraging and listening to new ideas, and creating an environment where these new ideas can flourish. Your participation in the FAC activities is vital for the production of data suitable for its intended use **and** may have an influence on the future path of INELA as a consensus standards development organization. All are encouraged to join INELA and to get involved! More information on the efforts of the INELA FAC may be found on the INELA web site (www.inela.org).

© Authored by Dawn D. Thomas, ASQ CQM, 2005. Thomas is a corporate Quality Assurance Manager for Professional Service Industries, Inc., a national engineering consulting firm, in Orlando, Florida and is certified by the American Society for Quality (ASQ) as a Quality Manager (CQM). She is the current Chairperson of the INELA Field Activities Committee and is a member of the INELA Board of Directors.

Development of a Comprehensive Quality Standard for Environmental Laboratory Accreditation

The Institute for National Environmental Laboratory Accreditation (INELA) is a not-for-profit research and scientific voluntary consensus organization whose mission is to help maintain and promote a system for the accreditation of entities directly involved in the generation of environmental data. To fully appreciate INELA's approach to standards development, one must be cognizant of pre-INELA standards development efforts, which although somewhat successful, were not without shortcomings.

The standards development process was not a true consensus process since private sector participants had limited input and did not have voting rights on the passage of the standards. The standards were focused on the classical perception of an environmental laboratory and did not address areas such as sampling or field measurements. The standards contained requirements for laboratories, government accrediting authorities, providers of proficiency test (PT) samples and their oversight bodies, and laboratory assessors. However, the requirements for each stakeholder group were dispersed throughout several of the chapters (e.g., the chapter on proficiency testing contained requirements for the laboratories, the PT providers, their oversight bodies, and the accrediting authorities that used the results of proficiency testing) making use of the standard quite challenging. The resulting standard lacked flexibility in its application, consisting of a set of interlinked chapters, none of which could be applied independently of the rest.

Growth of a national accreditation program was severely limited by this lack of flexibility. Further, the previous national accreditation standard was developed through incremental changes that were incorporated into the standards annually. Although the incremental approach was initially useful for identifying deficiencies and shortcomings through the standards' application and use, it represented a moving target, requiring users to continually adjust to new changes in the standards and accrediting authorities to continually modify their regulations for laboratory accreditation.

A Modular Approach

As a first step to adequately address the vulnerabilities of the pre-INELA national accreditation standards, INELA has initiated a process to establish utilitarian, stand-alone documents requiring infrequent modifications.

The new INELA standards will consist of modules, designed to serve as stand-alone documents, each meeting the specific needs of a particular user group. For example, one such module could describe the process by which an organization seeks to obtain and maintain accreditation. Another module could describe the requirements that proficiency test providers must meet. This change will obviate the current need for every stakeholder group to extract information from multiple chapters, where the requirements of other stakeholders are interwoven.

Consensus Standards Development

The second goal of the INELA standards development effort is to create standards in a true consensus environment, involving all interests and stakeholder groups. To this end, INELA has developed a document titled, *Policies Governing Standards Development*. These policies are posted on the INELA website (http://www.inela.org), and have been provided to the American National Standards Institute (ANSI) with an application for recognition by ANSI as an Accredited Standards Developer. According to OMB Circular A119, a voluntary consensus standards body must meet defined criteria, including balance of interest, lack of dominance, openness, and having an appeals process. As described in its policies, INELA achieves:

- **balance of interest** through Expert Committees with representation from every stakeholder group, all with full and equal voting rights;

- **lack of dominance** through an Expert Committee structure that prevents any single interest category from dominating any committee;

- **openness** through unrestricted membership, allowing all INELA members to comment and provide input on any proposed standard, and debating proposed standards publicly; and

- **consensus** and **due process** by providing the right to vote and register written comments on every proposed standard to every member, and to provide the opportunity for members to reconsider and change their votes. The INELA *Policies Governing Standards Development* also includes an appeals process whereby any member whose negative vote has been found non-persuasive may submit a written appeal for formal consideration.

INELA Committee Structure

INELA standards are developed by individual Expert Committees, which require balanced participation by affected stakeholders and lack of dominance by any single interest category. Each Expert Committee has a core of five to fifteen Committee Members and an unlimited number of Associate Committee Members, all having the right and opportunity to provide input on standard development and to vote on the standard before it can be released from the committee. Non-government members have the same rights as government members to provide input and vote on the standards. Any Committee Member may be the chair of an INELA Expert Committee. Committees also invite and consider input from any interested stakeholder regardless of INELA membership status.

There are seven Expert Committees that have been established to align with the existing chapters of the INELA Standard. As the current standards are restructured into a series of modules that will each address the needs of a particular group of stakeholders, the realignment and possible formation of new Expert Committees may be necessary. For example, a new committee may be formed to address the accreditation of critical agent testing, with its attendant requirements of national security consideration. The INELA Bylaws allow flexibility to expeditiously establish new committees, drawing upon the diverse expertise of its more than 500 members, and to

disband a committee whose objective is no longer aligned with the standards development goals of the organization.

The Expert Committees conduct most of their business through teleconferencing and two public meetings per year. The committees actively seek stakeholder input through processes that have been developed and described in the *Policies Governing Standards Development*.

Process for Standards Development

Standards development follows a systematic process that allows review and input by all affected stakeholders prior to voting on the final standards. The step-by-step process, as outlined in the *Policies Governing Standard Development*, is summarized as follows:

- Each Expert Committee develops a working draft standard, and votes to release this document for publication and public debate at the INELA semiannual meeting.

- Following the semiannual meeting, any INELA member or member of the public may submit comments to the Expert Committee, which considers the comments and makes appropriate modifications, to produce a Draft Interim Standard (DIS).

- All Committee Members and Associate Committee Members are afforded the opportunity to vote on the DIS, and to accompany their votes with written comments. The DIS becomes the Interim Standard if it receives a two-thirds favorable vote by the Committee Members and if all comments are resolved.

- The Interim Standard, together with the written comments, is debated publicly at the next INELA semiannual meeting. Further changes to the Interim Standard are not permitted.

- All INELA members then have the opportunity to vote on the Interim Standard. If the vote is favorable, the Interim Standard becomes the INELA Final Standard, which is then available for use by any standards adoption organization, such as the National Environmental Laboratory Accreditation Conference (NELAC).

The INELA standards development process has a distinct advantage over previous processes because the step-wise process allows ample time for members to consider the changes prior to voting, and affords members an additional opportunity to submit comments and have those comments resolved before the standard is accepted. This process will inevitably lead to a standard with fewer flaws, and hence, reduce the frequency by which further modifications to the standard will be required. INELA encourages anyone who is interested in this effort to join our organization and get involved in these efforts.

Advanced Tracking of Laboratory PT Performance and Certification Status with Integrated Electronic NELAC-Style Auditing Software

Tim Fitzpatrick, Laboratory Data Consultants (LDC), 7750 El Camino Real, Suite 2L, Carlsbad, CA 92009

Abstract

Regulatory Agencies, Laboratory Certification Bodies and QA oversight personnel are faced with the daunting task of tracking and monitoring the certification status of numerous laboratories for multiple fields of testing (FOT). The difficulty of accurately tracking certification status is further compounded by the requirement that laboratories demonstrate on-going competency by performing satisfactorily in annual or semi-annual performance testing (PT) studies for each accredited FOT. Some accrediting authorities must track performance results for two or more studies on hundreds of FOTs performed by hundreds of laboratories each year. Certification for each particular FOT is also contingent on satisfactory results from routine on-site audits.

Laboratory Data Consultants, Inc. (LDC) has developed a Microsoft ACCESS based software suite which automates the process of keeping critical PT and certification information up-to-date and accurate. PT study results available in electronic format from approved PT vendors can be easily uploaded into the central database. Once uploaded, the software evaluates the data and alerts the user when a particular lab is in warning or failure status for a particular FOT. In addition, the software tracks the certification status for all laboratories in the database from the initial application for certification through approval and the renewal process. Each laboratory's status for a particular FOT is available at the click of a button. The program also generates all required form letters for communication with the labs regarding PT and certification status.

The software suite also includes LDC's automated audit software, a Tablet-PC based program that guides the user step-by-step through a NELAC-style audit and automatically prepares the audit documents, summary table of findings and the audit report. The software comes loaded with the latest NELAC audit checklist, but can easily be configured to accept any state or client specific audit checklist.

This presentation shows how LDC's software suite makes the accreditation process more consistent, technically sound, and cost effective, resulting in more accurate and reliable data for the regulatory authority, the labs and the public.

Introduction

The amount of information that Regulatory Agencies, Laboratory Certification Bodies and QA oversight personnel have to keep up with on a daily basis is overwhelming. Laboratory certification authorities, for example must keep track of the overall status of possibly hundreds of labs as well as at least the three most recent results from PT studies for each accredited field of Testing (FOT) and how those sets of data interact with and affect one another. In addition to tracking the progress of a laboratory through the initial accreditation application and demonstration-of-capability, laboratories, in most instances, must successfully complete an on-site audit both initially and again periodically to maintain their certification.

Keeping track of all this information has been a time-consuming and cumbersome process. Some agencies have developed database programs in-house and some have purchased third party programs designed to manage all this information.

Laboratory Data Consultants (LDC) has developed the first bundled software suite to manage all this information in one location. LDC's Microsoft Access-based automated audit software program has been on the market for close to one year. LDC is currently developing additional modules that manage and incorporate information regarding the certification process and PT results management into one bundled application.

Certification Module

Upon initial application, the user enters information for the applicant laboratory and determines if the application type is for Primary NELAC certification, NON-NELAC certification, or secondary NELAC certification. The user is then prompted to enter information and verify and document completion of each critical step in the certification process.

The overall status of the application is automatically updated as each portion of the process is completed. The user is reminded by email of each of the critical dates in the certification process and of the dates when PT study data are due from each of the labs.

The status for each field-of-testing (FOT) is linked to the PT status information and is updated as the results history table is completed for each FOT.

Pre-defined queries provide information regarding certification status. For example, the user can run a report on all types of certifications held by a particular laboratory or get a complete listing of methods for which a particular laboratory holds certification. The user could also query the database for all laboratories which are in a certain status, e.g., all laboratories pending renewal in a certain timeframe.

Correspondence with the labs is automated by use of form communication letters such as notification that the laboratory's application has been received, when corrective action reports are required for failing PT data, notices of renewal applications, etc. The user just selects the type of letter required and fills in the required data for that specific template. The letter is then saved and attached to the laboratory's record.

Performance Testing (PT) Module

The PT tracking module imports electronic data deliverables (EDD) from PT providers directly into the program. The EDD import specs are taken from the USEPA National Standards for Water Proficiency Testing Studies Criteria Document, December 30, 1998. This is a format that all providers are familiar with and can currently produce.

After uploading the data, the user can then view the data in several different report formats, such as results grouped by laboratory, by state, by overall study, by date range, by result status (i.e., acceptable, not acceptable, check for error), etc. The user may also view a historical report summary showing the three most recent results for a particular FOT by laboratory along with a column evaluating the current status of the FOT based on the history. Possible results for status would be acceptable, warning (one failure in the two most recent results) suspended (two unacceptable in three most recent) or revoked.

After uploading the PT data, the user is notified of any changes in the status of a FOT and that status is updated in the certification module of the program as well.

When preparing for an on-site audit, the user can run summary reports of PT status to be imported into and shared with the automated auditing module of the software suite.

Communication templates are available in this module as well, such as a form letter for requesting corrective action for unacceptable results.

Auditing Software Module

No certification process would be complete without an on-site audit. LDC's NELAC auditing software is an integral portion of the software suite. This portion of the software guides the auditor through an audit using the 2002 NELAC checklist as the auditing standard by default. If the auditing agency wishes to conduct their audit using different standards, it is a relatively simple matter to upload other electronic checklists to use as the standard(s). The auditor simply picks which set of checklists to use when preparing the audit-specific template document.

The auditing software has two separate components, referred to as the Central

database portion and the briefcase database potion. The auditor does all the preparation work in the central database while connected to the network server where the database is housed. The first step involves importing relevant documents (QMs, SOPs, previous audits, etc.) into the central database. These documents are either supplied in electronic format by the lab or scanned and converted to pdf files for electronic storage and retrieval.

The auditor tags the documents and references to be imported and can hyperlink these documents to questions in the checklist. The auditor can then open any of the references during the course of the audit by simply clicking on the embedded hyperlink. The auditor next chooses the methods to be audited from the method and analyte database included with the program. This is essentially the listing of NELAC analytes and methods. If the user is auditing methods or analytes for which no code exists in the NELAC database, a wizard-style interface is available to create custom method and analyte codes.

Once the audit template is populated with the appropriate checklists, relevant documents and methods, the information is bundled and downloaded to an electronic "briefcase" database which is then transferred to a tablet pc and taken to the on-site audit. In the case of multiple auditor scenarios, the template may be downloaded to more than one tablet pc. The separate portions of the audit will be collected by the lead auditor at the conclusion of the audit and the software will combine the portions into one audit on the lead auditor's briefcase pc.

Although the auditor may use any type of portable pc as the platform for the briefcase, the tablet pc is the preferred platform primarily due to the ease of data entry compared with a standard laptop. The screen on a tablet pc may be flipped over so that the pc is the size and shape of a standard clipboard and data can be entered with a stylus by using the embedded handwriting recognition software in the tablet pc. This saves the step of having to print out the checklist(s), fill it out in hardcopy format and re-enter the results electronically at a later time.

Questions can be added to any checklist, either during preparation in the central database or during the audit in the briefcase. Additionally, documents such as scans of run logs, chromatograms, and even digital pictures can be uploaded to the briefcase during the course of the audit.

A tabular summary of findings and an audit assessment report are generated based on responses to questions in the audit checklist and the opening meeting worksheet. The summary of findings is then immediately available for distribution to lab staff at the closing meeting. The tabular summary can be exported in MS Excel format as well. This allows the laboratory to address each deficiency in the table electronically. The laboratory's response can then be uploaded directly into the summary table.

The assessment report is prepared from a template. As the auditor answers

questions on the checklists and in other locations, that information is inserted into placeholders in the report template. Once the placeholders are filled, the report is fully editable by the user. This allows the auditor to add free-form conclusions or expound on deficiencies listed in the tabular summary of findings and recommendations.

The report is sent to the lab in both hardcopy and electronic formats. The laboratory's response is then imported and a letter of acceptance or further action required is generated from a template. This process is repeated as necessary until resolution of all issues is complete. Once completed, the briefcase database is then uploaded back to the central database for archiving, virtually eliminating the need for paper filing and storage.

In addition to the current features, plans are in place to add an additional module which would guide the user through a document review audit (i.e., a paper audit with no on-site visit) as well as a training tracking module which QA personnel could use to centralize and retrieve employee training records and requirements.

Conclusion

There has always been a need for software products that can facilitate assembling and organizing the huge amounts of information associated with PT testing, laboratory certification and on-site assessment. LDC's suite of software products integrates all this functionality into one software package, virtually eliminating the manual tracking of information associated with these tasks. This product streamlines the process and eliminates potential sources of human error in assessing the data.

Attachments

A series of screenshots follows this section showing details of some of the key screens and features of the software.

Figure 1 – Central Database Main Screen

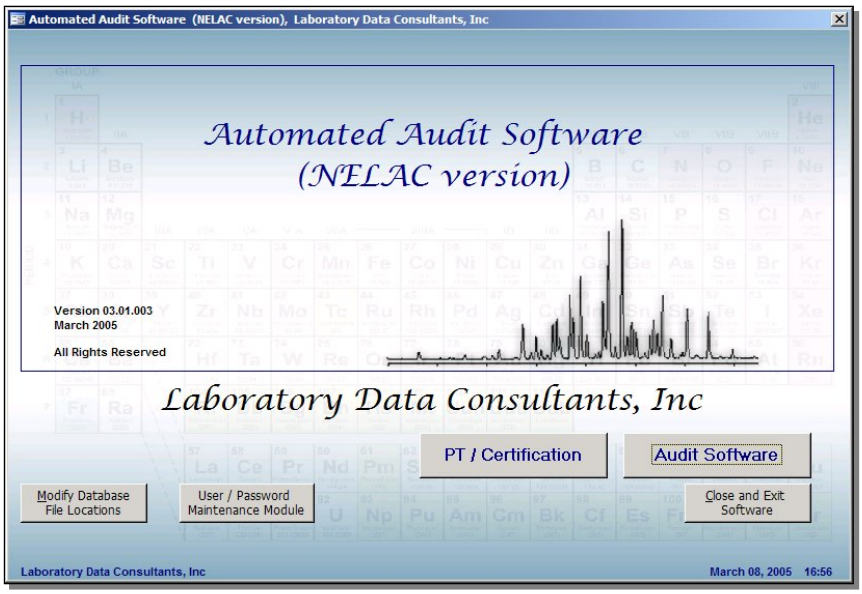


Figure 1 shows the main screen of the Central Database. From here, the user will choose to enter either the PT or certification tracking portion of the software or the Central database portion of the auditing software. The PT and certification tracking modules are housed only in the Central database.

Figure 2 - Briefcase

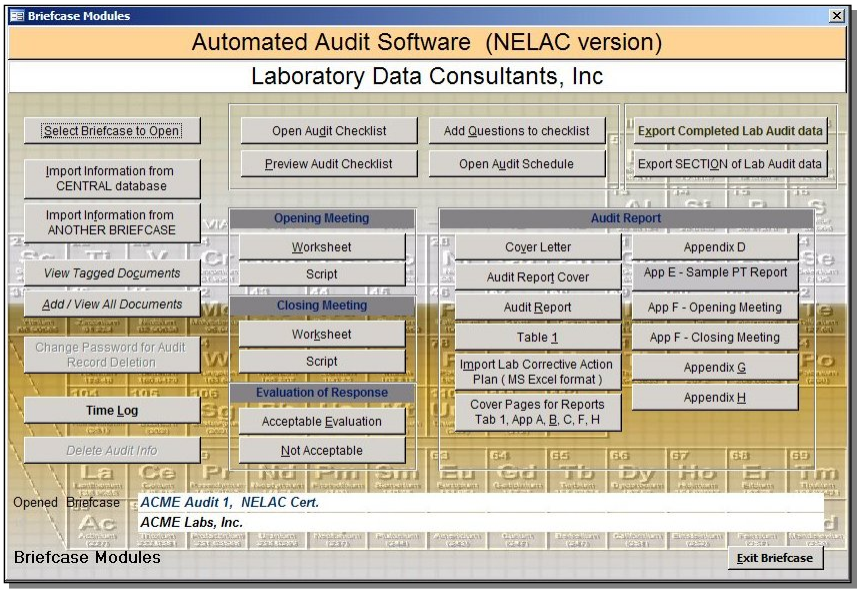


Figure 2 shows the main screen of the Audit Briefcase database. This is where the on-site audit is conducted and reports are prepared.

Figure 3 – Certification Main Screen

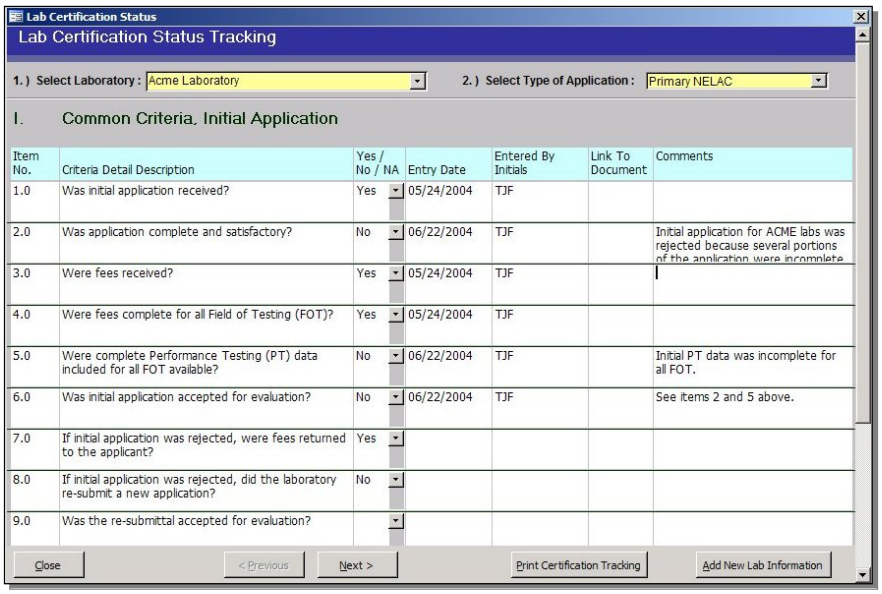


Figure 3 shows the main screen of the Certification module. This is where the user will track the progress and status of a particular laboratory's initial and renewal applications.

Figure 4 – PT Results View

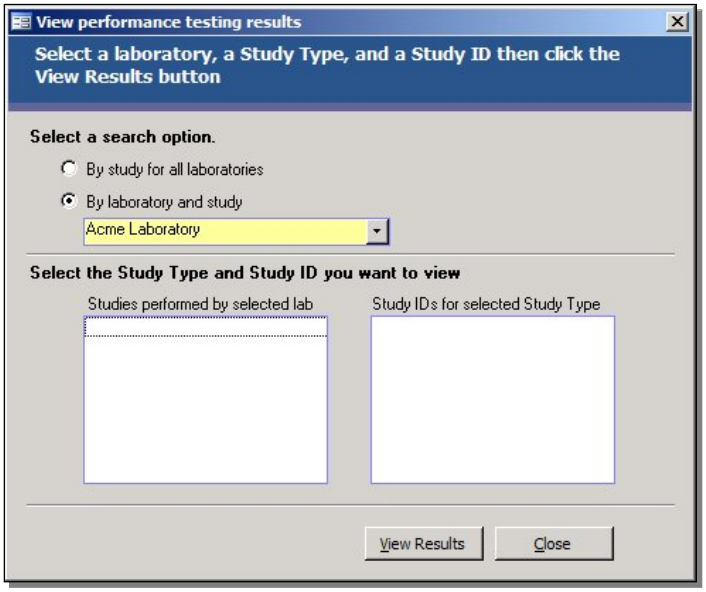


Figure 4 shows the interface where the user chooses the parameters by which to view the PT data. A separate historical summary report is also available for viewing as well.

Figure 5 – Audit Checklist

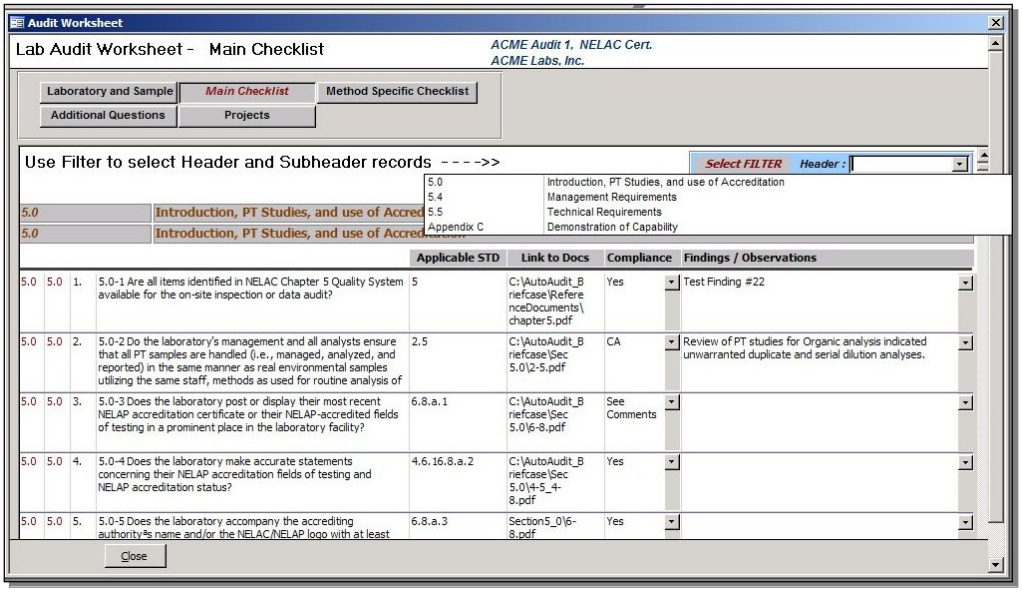


Figure 5 shows a portion of the main audit checklist. This is the 2002 NELAC checklist. Separate sections of the list are opened by clicking on the appropriate button and selecting the sub-sections of the checklist from the pull-down menus. The method-specific checklists (Appendix D1-D6) are accessed in a similar manner

Figure 6 – Tabular Summary

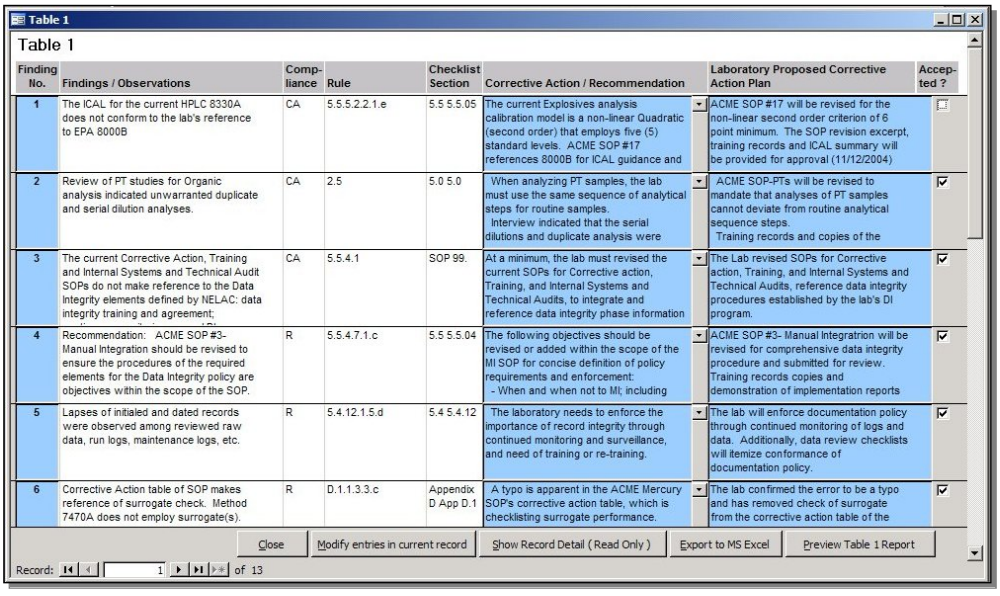


Figure 6 shows a completed tabular summary. Note that the lab's corrective actions have been imported and the first line item was not considered acceptable by the auditor. This item can then be re-exported and sent back to the lab for a second effort.

Figure 7 – Assessment Report

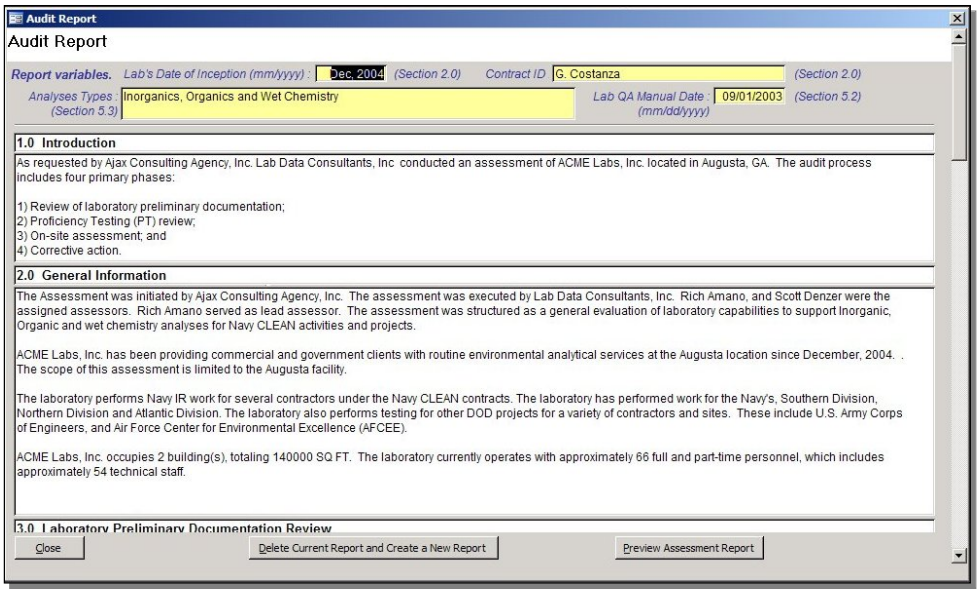


Figure 7 shows the Assessment Report template. The report is prepared automatically based on entries made in the checklists. The auditor may add freeform conclusions and make modifications as necessary to the text of the template.

Figure 8 – Multiple Auditors

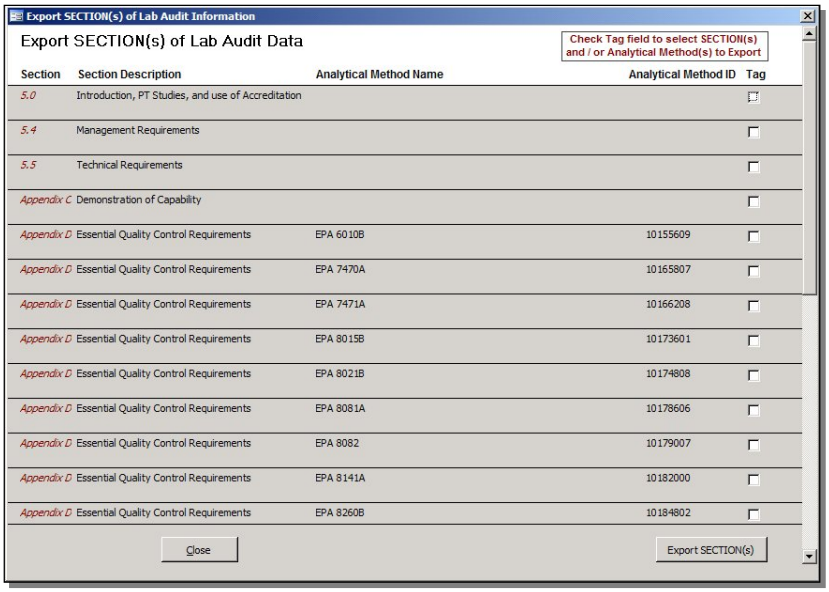


Figure 8 shows the screen where the supporting auditors go to export their portion of the completed audit. The lead auditor then imports the files prepared in this manner and collects them all in one central briefcase file.

Abstract: EnPPA Quality Management Metrics

Pat Mundy, USEPA/OEI/Quality Staff

January 18, 2005

Environmental Performance Partnership Agreements (EnPPAs) are incorporating more specific metrics to assist with the oversight of grant funds from USEPA Regions and Programs to States. During the current planning cycle, a specific program of the Wisconsin Department of Natural Resources (WDNR) was offered as a pilot to develop metrics that specifically address quality management concerns of Region 5 and the Great Lakes National Program Office (GLNPO). The pilot effort was cooperatively planned by the presenter (on detail at WDNR from OEI's Quality Staff), the EnPPA coordinators, and QA staff of WDNR, Region 5 and GLNPO. The process for developing the metrics, implementation progress and results to date will be presented.

Marcus E. Kantz
US EPA Quality Staff and Region 2 732-321-6690 kantz.marcus@epa.gov
Talk Presented at the EPA National QA Conference
April 13, 2005

How Good is "How Good is..."? (Measuring Quality Assurance)

The Quality Assurance (QA) community is in business to lead the environmental measurement community down the golden path toward collection of data that are appropriate for their intended use. In order to do this the QA community has itself devised and implemented a myriad of measurement techniques for assessing the quality of those environmental data. However, one job at which the QA community has not excelled is the measurement of its own effectiveness. We - at least those of us in EPA - have not yet figured out a satisfactory means for measuring and judging the effectiveness of our own QA programs. The purpose of this presentation is to discuss EPA's current initiative to improve the measurement of its QA program's effectiveness, and to offer some thoughts on expansion of this process to other organizations as well.

In the QA community, it's our job to assess how well people are doing, to measure how good data are, and to compare that to the levels prescribed in the underlying plans. We pride ourselves on being able to provide and oversee systems that answer the question: is it good enough? But now, when it comes to ourselves, and our own work, we don't have a very effective system for demonstrating how good we are at showing how good the data are. How good is "How good is"? Right now, we don't really know.

EPA's national program utilizes the Quality Assurance Annual Report and Work Plan (QAARWP) as the primary means for soliciting information from all of the Regions, Program Offices, and Laboratories on the status of their respective Quality Assurance Programs. A QAARWP template has been sent to each office, presenting a series of questions designed to encompass all that is important in their QA programs. This, in itself, is not a bad idea, except that the questions in the QAARWP template have not kept pace with the changing priorities in the Agency and with the changing functions within the QA programs. Many people have found that the data gathering and reporting process has not helped them to understand their own situation any better. Even worse, the questions reflect the QA program's own image in the data mirror, while largely ignoring the interaction of the QA program with the rest of the Agency, the very people that the QA program is designed to serve. The Quality Staff, the Agency's QA Headquarters office, has also found it very difficult to use the results provided in the QAARWPs to provide useful feedback to the Agency's offices on the status and usefulness of their QA programs or how they might improve them.

This spring and summer, the Quality Staff is undertaking a major initiative to re-invent the QAARWP, getting rid of lines of questioning that did not measure important aspects of the Agency's Quality System or its implementation, and adding questions that do. We are looking for ways to measure things that really matter to the QA community and to the rest of the Agency, our customers. We are starting to bring those customers - the Program Offices, the Labs, and the Regions - into the QAARWP development process to make sure that the new QAARWP will at least begin to measure aspects of the Quality Program that are meaningful to them. We understand and acknowledge that this will all be a difficult task, but, fortunately, the QA program comes to our rescue. We already have a tried and true, sound procedure for designing a data gathering exercise. Granted, we usually gather data about the water, the air, or the soil, but why shouldn't it work on data about the data quality assurers? The natural way for QA people to develop the new QAARWP is with a Quality Assurance Project Plan (QAPP).

In order to incorporate the views of the QA and Program communities, we plan to form a QAPP development Team composed of both, and charging the Team with the task of developing a set of questions for this year's QAARWP that begin to address the real issues facing the Agency's environmental data programs. We will ask the Team to start by addressing the Agency's most significant data gathering activities, the largest grants and contracts, the largest in-house projects, the major data bases. This way, the reporting burden on the Programs, Regions and Labs will be focused on those programs and projects that draw the most resources, produce the most important results, and expose the Agency to the greatest vulnerabilities.

While this is probably the most important area to be measured in the 2005 QAARWP, there is still considerable potential value in measuring other aspects of the Agency's Quality program. It is convenient to utilize a Framework that looks at the broad range of potential metrics, from Inputs through Outcomes. The attached chart describes the Framework in detail. What is really needed here are the metric categories, defined in the Quality Program context as follows:

Inputs:	Resources devoted to the Agency's QA activities
Activities:	QA work performed
Outputs:	Products directly resulting from the QA activities
Interim Outcomes:	Effects of QA activities on program objectives, which in turn facilitate progress toward the Agency's environmental goals
Outcomes:	Effect of QA outputs and interim outcomes on the accomplishment of Agency goals and/or statutory requirements

Ideally, the QAARWP would collect data on **Outcomes**, showing the degree to which the Quality Program actually affects the accomplishment of Agency goals, like clean air and drinkable water. However, it is very unlikely that a way could be found to sort out the contributions of the Quality Program from all of the other contributing factors. Thus, at least for the present, there will be no serious attempt to incorporate Outcomes metrics in the QAARWP.

As a surrogate, measuring Interim Outcomes is at least a possibility. For Interim Outcomes, the Quality Program goal is to improve Agency environmental decision-making and to improve Agency environmental data, information and products. In order to begin to measure these, reasonable metric goals would be the improvement in Agency decisions and information products, and a reduction in vulnerability to outside challenge due to QA. It may not be easy to measure improvements in Agency decisions, as the only source of information may be the managers' subjective opinions. However, it may be possible to measure improvements in data and information products and the reduction in vulnerability indirectly through the reduction in successful outside challenges. This may not be incorporated in the 2005 QAARWP, but shows promise for future years.

This now brings us to discussion of the three metric categories that can be measured directly and that still have the potential to provide clear insight into the effectiveness, efficiency, and usefulness of the Quality Program: Inputs, Activities, and Outputs. There is no question that these categories sound ominously like our old friends, the Bean family, and they can be. Thus, before we actually discuss the three categories and what sorts of metrics we might consider or recommend, we need to make clear the difference between a useful metric and a Counting Of the Beans (COB, what's left of the corn, after everything of value has been stripped off).

To coin a phrase, a Bean is the reporting of something that can be measured, estimated, or guessed at, and may even seem interesting, but the reporting of which adds little to nothing to our understanding or appreciation of the world. A very nice QA Bean would be: How many QAPPs did you receive? It sounds useful, at least for judging how busy the QA staff is. But without knowing if they actually tried to review the QAPPs, or even whether the State has already reviewed them, or whether they arrived after the projects were completed, or whether they were for projects the organization paid for, or cared about, or even knew about, the metric is just a COB, and the response is just a Bean.

A useful metric, on the other stalk of corn, is an EAR, Entirely Appropriate and Relevant. The data to be collected by the proposed metric must themselves be understandable and must contribute to an understanding of the effectiveness, efficiency, or value of the Quality Program. In comparison with our earlier COB, a nice fresh EAR might be: How many of your 10 largest environmental data contracts did **not** have their program and project level QA documentation (QMPs and QAPPs) properly approved? We know immediately what it means, it gets at very important work, and it will reveal whether that work was covered under the Quality System, including both organizational control (the QMP) and systematic planning (the QAPP).

Therefore, we will tell our QAPP development Team that the purpose of this data gathering project is to prepare a set of EARs, without any COBs. And now, finally, we are really ready to talk a little bit about the actual types of metrics that we will suggest they consider.

To start, we of course will tell them to aim to include as many **Interim Outcome** metrics as possible. We really need to understand the effect that the Quality Program has on the end products of the Agency's work. We will suggest that the Team investigate issues such as:

> What fraction of the Agency's major decisions were based on data projects that had conducted and documented systematic planning and assessment?
> What fraction of the Agency's significant information products and data bases had conducted and documented systematic planning and assessment?
> What percentage of challenges of properly QA'ed information products were successful, vs. the percentage for products that were not properly QA'ed.

These three ideas investigate different aspects of whether we are implementing our QA resources where they matter most, according to our plans, in open verifiable ways, and with a degree of success. In future QAARWPs, we will also try to discern whether the Agency's decisions themselves are improving over time.

After Interim Outcomes, we'll tell them to include a smattering of Inputs, Activities, and Outputs, because we need to be able to assess the relevance of our Process, not just our results. Are we doing things that make sense, that contribute to attaining our goals, and that are documented?

For **Outputs**, we will suggest that the Team consider issues like:

> What percentages of the required QMPs, QAPPs, and corrective actions following major audits are in place?
>
> The purpose of these questions would be to document the actual implementation of the organization's Quality System
>
> (Note, one could simply ask: Are you implementing your Quality System? But that would be a Bean, and gets at another un-tasty characteristic of Beans, they are often undocumented and unverifiable.)

For **Activities**, we will suggest inquiries such as:

> What percentage of your major QA commitments did you achieve, such as: training given, guidance delivered, systematic planning facilitated (QAPP reviews) assessments conducted, feedback provided...?
>
> These questions would document the relative progress of the organization toward achieving its QA commitments to the organization's staff and programs.

Finally, for **Inputs**, the goal of the metrics would be to compare the organizational allocation to the Quality Program to its stated levels of commitment. Questions might be like:

> What percentage of your committed annual budgets for QA training, guidance, assessments, and assistance was actually expended?

The Agency's Quality Assurance community is committed to maintaining and implementing a Quality Program aimed at assisting the Agency in achieving its mission by providing an infrastructure for well planned, implemented, assessed, and documented environmental data operations. A vital element in that infrastructure is relevant monitoring of its status, progress, successes, and problems, and presenting the information to those who can use it. The QAARWP can and should be a cornerstone of that infrastructure. Our goal is to make it so. We will also suggest that our partners in the States, Tribes and elsewhere consider incorporating the QAARWP concepts into their own Quality tracking and reporting process.

As we develop and roll out our improved QAARWP this fall, we hope that the Agency's program and QA communities will endorse and utilize it for their combined benefit. Quality Staff will analyze the submitted results and report back to the respondents with comments and suggestions. This will also serve as a valuable assessment of the QAARWP itself, by determining whether it truly solicited information that proved useful for assessing the quality of the respondents' Quality Programs.

We anticipate that the 2005 QAARWP will still be a work in progress. We will make efforts to ask useful questions and to provide useful comments to the respondents. Of course we will also hope and expect to receive a lot of constructive feedback from all who report through the QAARWP and try to utilize the data that come from it, allowing us to do a better job in future years. Only then will we begin to understand how good is 'how good is?'

Framework for EPA QA Performance Metrics Covering Quality Staff, Regions and National Programs Proposed Metrics to Include Inputs through Interim Outcomes				
Metric Categories	**Definitions**	**QA Program Goals**	**Metric Goals**	**Prototype Metrics**
Inputs	Resources devoted to QA of programs	Expend QA resources properly	Percent of planned QA expenditures spent	Percent of budget plan allocations actually expended for major QA program areas (e.g., training, guidance, assessments)
Activities	QA work performed	Achieve QA program commitments and objectives	Relative progress toward QA program commitments	percent of major QA program commitment levels achieved (e.g., training, guidance, assessments)
Outputs	Products directly resulting from QA activities	Implement QA requirements & guidelines effectively; promote understanding and reporting of data quality	Frequency of: conformance to program and project QA requirements; identification and addressing QA problems	Percent of required QMPs and QAPPs in place; percent of required corrective action plans in place as a result of QA activities?
Interim Outcomes	Effect of QA activities on program objectives, which in turn facilitate progress toward Agency environmental goals.	Improve Agency environmental decision-making; Improve Agency environmental data, information and products	Improvement to Agency decisions, products, and reduction in vulnerability to outside challenge due to QA.	Percent of decisions reported to be improved; Percent of significant information products and databases that have conducted and documented systematic planning and assessment; percent of successful challenges of products properly QA'ed vs. percent for non-QA'ed products.

Outcomes	Effect of QA outputs and interim outcomes on accomplishment of Agency goals and/or statutory requirements	Improve Agency environmental results due to effective QA implementation.	Improvement in Agency implementation of statutory and regulatory programs due to QA	How much did discharges decrease and/or ambient conditions improve as a result of QA activities? [Likely unmeasurable now]
Impacts	Effect of QA program on the degree to which broad strategic goals or objectives are achieved	Improve Agency's achievement of strategic goals or objectives due to the QA program.	Effect of QA activities on human health and the environment	How much human health and/or ecological improvement resulted from QA activities? [Likely unmeasurable now]

PERFORMANCE-BASED MANAGEMENT

Mr. James Gonzales, Ms Erica Becvar, Dr Javier Santillan
HQ AFCEE / TDE
3300 Sidney Brooks
Brooks City-Base, TX 78235-5112
Telephone: 210-536-4366
E-mail: Javier.Santillan@brooks.af.mil

The Deputy Secretary of the Air Force has expressed her concern at the apparent slow pace with which the Air Force environmental liabilities are being reduced. AFCEE has been aware of this problem, and has been involved in the development and implementation of approaches to accelerate cleanup of sites. The optimization programs implemented have had some successes, but also have identified root causes for the deficiencies, and strategies required to attain success.

AFCEE needed to develop a holistic approach to accelerate the installations' environmental restoration program. This approach had to manage the uncertainties inherent in the restoration of highly heterogeneous sites. The solution is Performance-Base Management (PBM) of the Environmental Restoration Program. PBM incorporates the use of systematic planning, dynamic workplans, implementation of decision logic in all plans and documents, real-time data collection and workplan review, exit strategy, implementation of optimized operation of restoration process, and procurement of all services through performance-based contracts (PBC).

A general description of PBM will be presented followed by a more detailed presentation on the components and principles of decision logic, systematic planning, dynamic workplans, and exit strategy.

A Quality Assurance Project Plan Guidance for Wetlands Projects

David R. Taylor, Mary Butterwick, Mark Kutnink,
Cheryl McGovern, and Eugenia McNaughton

US Environmental Protection Agency Region 9

The US EPA Region 9 Quality Assurance (QA) Office in conjunction with representatives from the Region 9 Water Division have developed a guidance document to be used by grantees who are performing environmental measurements related to wetlands. Although the elements which would be present in a Quality Assurance Project Plan (QAPP) prepared using the familiar EPA G-5 Guidance (Guidance for the Preparation of Quality Assurance Project Plans), have been retained, the guidance has been reorganized to better meet the needs of potential users. It is designed to reflect the emphasis on field measurements and observations common to most wetlands projects.

1.0 Introduction

For the purpose of the QAPP guidance, wetlands projects are defined as jurisdictional wetlands and non-wetlands aquatic resources (e.g., tributary streams, coral reefs, mud flats, etc.) and adjacent riparian vegetation. Wetlands projects are becoming more prevalent among those funded by EPA, at least in Region 9, as the role of wetlands in assuring the health of the nation's waters, and their role in such areas as flood control become better understood. However, Region 9 Water Division staff found that EPA's G-5 Quality Assurance Project Plan (QAPP) presented a number of challenges to grantees who potentially wished to use it to prepare a QAPP. Among these were the confusion associated with trying to understand the redundant sections in the guidance, the presence of sections which seemed to have little relevance to the types of projects they might conduct, the difficulty in defining data quality objectives (DQOs) and Measurement Quality Objectives (MQOs) for these types of projects, and the problems in adapting a document oriented toward fixed laboratory analyses to what were often, but not always exclusively, field based operations. In summary, the existing guidance was not optimal for use in describing the quality systems associated with wetlands type projects.

The result of these discussions was the creation of a small workgroup consisting of Water Division and Quality Assurance Office Staff. This workgroup reformatted the existing guidance and added information and examples which it was felt would be more useful and relevant to these types of projects. This new guidance document is described below.

2.0 Organization of Document

Wetland QAPPs prepared using this new guidance, while retaining most of the elements found in G-5, will organizationally be quite different. First, rather than 24 there are now 9 Sections. These are:

1.0 Project Description
2.0 Project Data Quality Objectives
3.0 Field Study Design/Measurement Protocol(s)
4.0 Field Preparation and Documentation
5.0 Quality Control for Samples Collected for Off-Site Analysis
6.0 Field Sample Collection Protocols for Off-Site Analyses
7.0 Laboratory Analyses and Selection
8.0 Sample Shipment to Off-Site Laboratory
9.0 References

The QAPP is organized so that a project which consists solely of field measurements and observations activities, a common situation for these types of projects, will require completion of the first four sections (plus references), whereas a project with both field and laboratory components would require completion of all sections. A more conventional water quality monitoring project, one without a significant field component, would not be expected to use this guidance, but instead would rely on G-5.

The types of information which would be captured in each section is described below. The discussion will focus on the first four sections, as guidance relevant to the collection of samples for off-site analyses is well established and in wide-spread use.

3.0 Document Contents

3.1 Project Description

This section is relatively self explanatory and consists of six sections. The information is similar to that requested in a standard QAPP, but has the plan writer focus mainly on wetlands examples and the factors which might influence the success of this type of project. Thus, “Project Purpose” orients the writer toward assessment, restoration planning, seasonal limitations, etc., factors relevant to wetlands projects. The “Project Area Description” section requests information on physical structure, hydrogeology, and biological aspects of the area to be studied, rather than sites of chemical spills, buildings, etc.. The section requests maps, both area scale and project scale.

The next two sections on “Responsible Agency and Participating Organizations,” and “Project Organization and Responsibilities” are both similar in scope to a standard QAPP, but the section following, “Permit Requirements for Collection of Environmental Measures” is not usually explicitly called out in a conventional QAPP. The nature of wetlands projects sometimes requires the grantee to obtain permits before collecting specimens for in field or later identification, and thus this needs to be included as part of the planning process.

Finally, the last section, "History, Previous Studies, Regulatory Involvement" emphasizes primarily discussions of previous studies. Most wetlands projects have not been part of actions taken by regulatory agencies as would be the case with more familiar forms of environmental problems, and often the traditional history of environmental problems found in many QAPPs is not relevant, since development, alien species, and water issues often represent a greater threat to wetlands than pollution.

3.2 Project Data Quality Objectives

It is in this section that the guidance attempts to offer suggestions to wetlands project QAPP writers to overcome challenges presented by existing guidance. To accomplish this four general sections are provided to assist in defining data quality objectives and decisions to be made with the data. The first section asks the writer to define questions to be answered, decisions to be made, and hypotheses to be tested. The guidance then provides examples within four areas which are commonly related to wetlands work. For the purpose of the guidance these include the categories of, "Inventory," "Assessment," "Trend Monitoring," and "Project Monitoring." A suggested decision action for inventory is, "If certain types of wetlands within the study area are found to support threatened and endangered species, then special land use zoning will be sought for these areas." For project monitoring, a suggested question is, "What is the change in functional condition as a result of your specific project?"

The second part of the DQO section requests the writer to "Identify the general categories of information needed to answer the questions." Again, examples are provided under the four categories of project described above. So for assessment, there are suggestions for a sediment loading and transport study. Such a plan might first list the parameters associated with accelerated sediment loading and then describe the measurements needed to assess how the wetlands have changed over time as a result of increases in sediment.

The third area in the DQO section requests the writer discuss, "What specific measurements and observations are needed to answer the questions?" Under trend monitoring, there are suggestions concerning what could be monitored, such as increased percentage of impervious surfaces in adjacent areas, increases in exotic species by frequency, type and magnitude of occurrence, etc.

The final section asks the question, "What criteria will be used to evaluate each of the measurements and observations?" The section provides provisions for both quantitative and qualitative criteria to be discussed. For example, under inventory, a standard classification system might be used which was prescriptive in how it classified different flora or fauna. This would lead to a conclusion concerning a percentage of dominant vegetation providing a correct wetland classification.
In cases where water quality measurements would be made, with samples collected for off-site chemical analyses, the defining of objectives, action levels or criteria, etc. was expected to be similar to a conventional QAPP, and is not discussed here.

Although this approach does not follow the Agency DQO process (however the Agency's DQO guidance document is referenced for further information), it should enable the effective defining

of project goals and decisions and the accompanying criteria to support achievement of those goals.

3.3 Data Quality Indicators for Field Activities

This section was presented in a more narrative form than might be found in a field sampling plan or QAPP involving chemical measurements. The DQIs of representativeness and comparability are addressed in traditional ways mainly through an emphasis on sampling design and the use of well established protocols and techniques. Completeness is also defined in conventional terms related to the number of valid observations and field data collections compared to those required for a successful project. Examples of precision are a significant component of field measurements, whether it is measuring flows, identifying macroinvertebrates, or conducting an inventory of alien species. Usually field teams have sufficient members that some sort of dual check is possible to confirm the counts or measurements made by another staff member. The options for accuracy assessment for many wetlands field measurements is limited, but in some cases there are activities which can be carried out (voucher specimens are discussed later). For example, ground truthing of remote sensing data, the use of established and recognized experts for confirmation of team activities, etc. Sensitivity relates more to chemical analyses, rather than the types of activities used in wetlands assessments, so this is mainly discussed in that context in the guidance.

3.4 Data Review

This section is straightforward, and the expectation is that each wetlands project will describe procedures to affirm and verify the results it generates.

3.5 Data Management

Emphasis is placed in this section on the transference of data from the field to whatever databases and records will be maintained after the project. The process is not that different from a conventional project, although the nature of the data being recorded is such that the procedures may require special forms or data record keeping.

3.6 Assessment Oversight

Assessment for wetlands projects is mainly discussed in terms of evaluation of the capabilities of off-site laboratories which might be involved in chemical, physical, microbiological, taxonomic, or other type activities.

3.7 Acquired or Secondary Data or Non-Direct Measurements

This section refers the QAPP writer to the G-5 guidance, as it was felt that evaluation of secondary data for these types of projects would not differ from that for other types of projects.

4.0 Field Study Design and Assessment Protocols

This section provides a number of suggestions on how to describe the activities which will be carried out under the project, since it is assumed that most such projects will rely heavily on field activities. To this end, a number of separate areas are discussed. This are divided into two areas, Physical and Chemical Characteristics and Biological and Habitat Characteristics. Contents addressed under these two sections include:

- Physical and Chemical Characteristics
  - Physical and Chemical Characteristics
  - Landscape/Watershed Scale Data
  - Geomorphology
  - Hydrology/hydraulics
  - Soil
  - Sediment
  - Water Quality
- Biological and Habitat Characteristics
  - Field Data
    - Vegetation
    - Habitat Assessment
    - Botanical Surveys
    - Faunal Surveys
  - Voucher Specimens for On or Off-Site Identification
    - Botanical
    - Faunal

For each of these areas, the guidance provides suggestions on approaches that might be used to define a quality system that will result in data of known quality. Details of these activities are outside the scope of this paper, but more information can be found in the guidance itself.

5.0 Field Preparation and Documentation

This section mainly provides guidance on the types of documentation which would be appropriate for wetlands activities. The section on "Field Preparation" is the one which differs the greatest from Agency guidance in that it requires a description of equipment and preparation activities relevant to wetlands projects. The other sections, on "Field Logbooks," "Field Data Sheets and Forms," "Photographs," "Documentation of Sample Collection,"and "Labeling of Sample Containers," do not differ significantly from standard QAPPs. Following this is a section with provisions for "Field Variances," which basically discusses contingencies for changes in the field.

6.0 Summary of Field Portion of the Guidance (Sections 1-4 of the guidance)

Provided that these sections were completed satisfactorily, this would constitute an acceptable document for a field-only wetlands project. The remaining sections of the guidance would not need to be addressed, aside from the last one on "References." In some ways the Reference section is especially important in a wetlands QAPP, since many of the field measurements and off-site voucher specimen work rely on the use of well accepted, standard protocols. To further aid the plan preparer, a number of these references are cited throughout the guidance.

7.0 Other Sections of the Guidance

Although it is expected that most grantees would use the field part of the guidance exclusively, there are situations when samples will be collected for off-site chemical laboratory analysis. The remaining sections of the guidance address these types of activities. This part of the guidance will be briefly described. There are four additional sections. These include:

Quality Control for Sample Collection for Off-Site Analyses
Field Sample Collection for Off-Site Analyses
Laboratory Analyses and Selection
Sample Shipment to the Off-Site Laboratory

The first of these sections requires a discussion of the types of quality control samples that will be collected and the measurement quality objectives (criteria) that will be used to evaluate the data. This section was intentionally moved from the traditional DQO section to keep the field work separate. However, DQOs (decisions and criteria upon which decisions will be made) would have been discussed in the earlier DQO section, if water quality monitoring were a part of the project. In some cases suggested language is provided to facilitate the preparation of the discussion and to provide the writer with the appropriate level of detail required.

The section requires discussion of sample collection activities as they relate to off-site work. The guidance prompts the writer to discuss the collection of all types of samples that might be relevant including water, soil, sediment, groundwater, foliage (for chemical analysis, not identification) and faunal (again, for chemical analyses, not identification). As with the quality control section, selected text is provided in some cases to give the grantee a better idea of the information required.

The next section requires a discussion of laboratory analyses, and also prompts for more information if less standard methods will be used, such as might be required for chemical analyses of botanical or faunal samples. There is also a requirement for a discussion of laboratory selection.
Finally, the guidance contains a section covering sample preservation and shipment, a section routinely found in most QAPPs.

8.0 Conclusions

It is hoped that the guidance described in this paper will be of benefit to those grantees and others who are required or desire to document their approach to quality assurance for wetlands projects. The guidance, with its strong emphasis on the field aspects of such projects, hopefully provides another tool in the toolbox for Regions, states, and tribes who might we working in the growing area of wetlands development. The guidance can be downloaded from the Region 9 web page, at http://www.epa.gov/region09/qa.

Military Munitions Response Program (MMRP)
Quality Plans

John Sikes, U.S. Army Engineering and Support Center, Huntsville

References

1. ANSI/ASQC E4-1994, Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs

2. DODI 4715.XX, Environmental Data Quality Assurance, Draft May 2003

3. DOD 4715.XX-M-1, DOD Quality Systems Manual for Environmental Laboratories, Final Version 2, July 2002

4. DOD 4715.XX-M-2, Uniform Federal Policy for Implementing Environmental Quality Systems

Currently available as IDQTF document; "Uniform Federal Policy for Implementing Environmental Quality Systems – Evaluating, Assessing, and Documenting Environmental Data Collection/Use and Technology Programs", Final Version 1, January 2003

5. DOD 4715.XX-M-3, Uniform Federal Policy for Quality Assurance Project Plans

Currently available as IDQTF document; "Uniform Federal Policy for Quality Assurance Project Plans – Evaluating, Assessing, and Documenting Environmental Data Collection and Use Programs", Final, July 2004

6. Example Quality Assurance Project Plan for Munitions and Explosives of Concern, Review Draft Version 1, September 2004

7. EPQ QA/R-2, EPA Requirements for Quality Management Plans, March 2001

8. Engineer Regulation (ER) 5-1-11, U.S. Army Corps of Engineers Business Process, 17 August 2001

9. USACE Procedure 2000 (PROC2000), Project Management Plan (PMP/Program Management Plan (PgMP) Development

10. USACE Reference 8005G (REF8005G), Project Management Plan (PMP)/Program Management Plan (PgMP) Content

11. USACE Reference 8008G (REF8008G), Quality Management Plan

12. Engineer Manual (EM) 1110-1-4009, Military Munitions Response Actions, Draft Final, 7 October 2004
13. Engineer Manual (EM) 200-1-2, Technical Project Planning (TPP) Process, 31 August 1998

14. Engineer Manual (EM) 1110-1-1200, Conceptual Site Models for Ordnance and Explosives (OE) and Hazardous, Toxic, and Radioactive Waste (HTRW) Projects, 3 February 2003

History

The ideas of Quality Control, Quality Assurance and Quality Plans continue to evolve as the Military Munitions Response Program, formerly called Ordnance and Explosives Response, continues to evolve and mature. Quality Assurance has been a part of each project from the very first project, from determining project goals to identifying minimum qualifications for selected disciplines. Project personnel didn't formally call anything they did "quality assurance" or even document their plans or their activities as such. Eventually the concept of quality assurance and the need to document plans and record actions taken has become clearer, but discussions among project disciplines and various agencies are still ongoing.

Standards, Regulations and Guidance

The references listed above provide numerous requirements and formats to determine and document quality plans by various names; Quality Assurance Project Plans, Quality Management Plans, Quality Control Plans, Quality Assurance Surveillance Plans. With each agency requiring similar documentation, albeit by different titles, project delivery teams (PDT's) can easily become frustrated when trying to apply different processes and meet the needs of stakeholders and regulators. This paper will demonstrate the confusion felt by PDT's while trying to meet everyone's needs and requirements. Below is a synopsis of the main sources typically quoted as requirements for development of quality plans.

ANSI/ASQC E4-1994
Specifications and Guidelines for Quality Systems for Environmental Data Collection and Environmental Technology Programs

This consensus national standard is intended as a guide and describes those elements determined to be effective in maintaining a reliable, consistent quality system regarding collection of environmental data. This standard discusses the need to have appropriate management systems in place to assure that resources are available, plans and activities are documented, and quality is assessed.

DODI 4715.XX (DRAFT)
Environmental Data Quality Assurance

This proposed DOD Instruction is intended to implement policy and assign responsibilities for implementing environmental quality systems into DOD activities and programs involving the collection, management, and use of environmental data. It also defines environmental data as any measurements or information that describe environmental processes, locations, or conditions; ecological or health effects and consequences; or the performance of environmental technology. This instruction also authorizes the publication of the DOD 4715 series of environmental quality guidance.

DOD 4715.XX-M-2
Uniform Federal Policy for Implementing Environmental Quality Systems

This guidance is currently under review. Current guidance on this topic is found in the Intergovernmental Data Quality Task Force (IDQTF) Uniform Federal Policy for Implementing Environmental Quality Systems – Evaluating, Assessing, and Documenting Environmental Data Collection/Use and Technology Programs. This policy provides implementation policies for federal agencies, therefore is more detailed than ANSI/ASQC E-4 or EPA's QA/R2. It is intended to apply to federal agencies whereas EPA QA/R2 applies only to organizations implementing programs for EPA.

It also represents consensus policy reflecting the views of the IDQTF. The development of this policy was enhanced by lessons learned from a broad set of stakeholders involved in managing quality across several federal agencies.

DOD 4715.XX-M-3
Uniform Federal Policy for Quality Assurance Project Plans

This policy provides for the development of project specific plans for the collection of environmental data and environmental technology programs. It was originally written for hazardous waste programs and federal agencies but provides good guidance for other environmental data collection activities as well.

Based on this document and recognizing the unique aspects of the MMRP process and data needs, the IDQTF has developed a companion example QAPP for a fictional MMRP project. It is currently being reviewed by task force members and their technical advisors.

Engineer Regulation (ER) 5-1-11
U.S. Army Corps of Engineers Business Process

This regulation establishes philosophy, policy, and guidelines to accomplish all work performed by the U.S. Army Corps of Engineers (USACE). It establishes the Project Management Business Process (PMBP) and applies to planning, development, and management of program as well as projects, and is used at all levels of the USACE. It

defines the project delivery team (PDT), which includes the customer, the project manager, technical experts, stakeholders, and regulators, as vital members.

This regulation defines 7 USACE Business Imperatives, one being *Plan for success and keep commitments.* Requirements for quality must be addressed during the planning phase. It is important to USACE to build trust with customers, regulators, and coworkers by clarifying expectations, keeping commitments, and ensuring projects are delivered as promised. To meet these objectives, all work will be managed under management plan. The PMP is intended to be the roadmap for quality project delivery. The PMP is an agreement between USACE and the customer that defines the customer's desired outcomes.

Among the required subplans in the PMP is the Quality Management Plan (QMP). The QMP is intended to reference project-specific quality assurance and quality control procedures appropriate to the size, complexity, and nature of the project. This QMP further defines customer quality objectives, addresses each major deliverable and quality assurance methods, and any other tasks or functions impacting quality project delivery.

Engineer Manual (EM) 1110-1-4009 (Draft Final)
Military Munitions Response Actions

This EM guides the USACE project team through the engineering design requirements that will be addressed while planning a MMRP project. Chapter 11 of the Draft Final version (final to be issued later summer, early fall 2005) describes the roles and responsibilities of the PDT with regard to development and implementation of a project specific Quality Assurance Surveillance Plan (QASP). The QASP directly corresponds to a contract's specified performance standards, is used to measure contractor performance and to ensure that the Government receives the quality of services call for under the contract and only pays for the acceptable levels of services received.

The QASP describes how government personnel will evaluate and assess contractor performance against established performance standards. The QASP documents performance metrics, when met, assure that project objectives are met. It documents specific roles and responsibilities of government team members including the types of surveillance methods to be conducted. The QASP will document the process to be used in instances of product or process non-compliance with established protocols.

General

All current USACE documents together fulfill the intent of the IDQTF QAPP guidance. Current quality plans/documentation required for USACE MMRP projects are:

1. Project Management Plan (PMP)
2. Quality Management Plan (QMP)
3. Quality Assurance Surveillance Plan (QASP)
4. Quality Control Plan (QCP)

The QAPP is intended to document the results of the systematic planning process. Within USACE, this process is called Technical Project Planning (TPP) which fulfills all requirements of the EPA systematic planning process, and ultimately results in the identification of data needs and development of data quality objectives (DQO's). A crosswalk between the 7 Step DQO process and TPP is provided in Appendix E of EM 200-1-2.

The results of the TPP process are typically documented using the worksheets provided in the TPP manual which are similar to those provided or suggested for use in the example MEC QAPP, reference 6. These DQO's are then used as the basis for development of contractor quality control and government quality assurance processes.

The contractual definitions of quality control (QC) and quality assurance (QA) state that the contractor is responsible for QC, and the government for QA. Contract requirements regarding the contractors QC Plan are not well defined, relying on the use of TPP to identify critical areas to be included in the contractor's QC Plan and the government's QASP.

Crosswalk of QAPP and USACE Requirements

Because of the similar information required of both the QAPP guidance and the USACE documents listed above, there is a need to show how they compare. The table below identifies example MEC QAPP information and the USACE document(s) where all or part of that information is typically found. Until USACE is directed to develop formal project specific QAPP's we will continue to develop the basic documents identified above. We do, however, suggest that the QAPP guidance be used as a basis for discussion among the USACE project team, stakeholders, and regulators. However, it would be unreasonable to expect that both sets of documentation be developed.

Example MEC QAPP Contents	TPP Process	PMP/ QMP	QASP	WP	QC Plan
Site description -Summary of previous work -Conceptual Site Model	X	X		X	
Project management, Organization, schedule, overview -QC program overview -Data management overview	X	X	X	X	X
Project objectives -Problem definition -Project scope -Quality objectives	X	X	X	X	X
Investigation design, field methods -Investigation methods -Sampling Plans -Geophysical prove-out -Anomaly ID and Excavation -Data management	X			X	

Assessment and oversight -Field operations -Documentation -Corrective action -Improvement -Identifying deficiencies and non-conformances			X		X
Data Review -Verification of Data -Data usability -Review of data	X		X	X	X

X- Denotes that all or parts of the QAPP contents are covered in the selected document.

Conclusion and Recommendation

One can easily see how the potential for confusion, disagreements and lengthy discussions revolving around quality plans and documents can get started. However, the project team, stakeholders and regulators should keep in mind that regardless of what the document is called, the basic requirements are the same, and the goal of every project is to make sound decisions that ultimately reduce the risk to the public and the environment presented by the presence of military munitions.

Until approval of the DOD guidance and implementing instructions, it is recommended that USACE require contractors to utilize IDQTF guidance, formats, and examples to develop their project specific QC and QA Plans. Stakeholders and regulators should understand the USACE requirements and assure that their quality needs are addressed.

NATTS QA System and Results of the QA Assessment

Dennis K. Mikel
Environmental Protection Agency - Office of Air Quality, Planning and Standards
Emission, Monitoring and Analysis Division
MD D243-02
Research Triangle Park, North Carolina, 27711
919-541-5511
mikel.dennisk@epa.gov

Abstract

In order to assess whether the population is at risk to Hazardous Air Pollutants (HAPs), the EPA in cooperation with the State/Local/Tribal (S/L/T) air pollution agencies are creating a long term national trend network. This paper will outline the Quality Assurance (QA) System that has been implemented for the National Air Toxics Trends Stations (NATTS) network. In addition, this paper will illustrate the results of the first and second set of Proficiency Tests(PT) for the laboratories that participated in the NATTS. This paper will discuss the results of those PT samples and illustrate how that data will influence the future testing program for the NATTS.

Introduction

There are currently 188 Hazardous Air Pollutants (HAPs), or Air Toxics (AT), regulated under the Clean Air Act (CAA) that have been associated with a wide variety of adverse health effects, including cancer, neurological, reproductive, developmental, as well as eco-system effects. These air toxics are emitted from multiple sources, including major stationary, area, and mobile sources, resulting in population exposure to these air toxics as they occur in the environment. While in some cases the public may be exposed to an individual HAP, people typically experience exposures to multiple HAPs from many sources. Exposures of concern result not only from the inhalation of these HAPs, but also, for some HAPs, from multi-pathway exposures to air emissions. For example, air emissions of mercury are deposited in water and people are exposed to mercury through their consumption of contaminated fish.

Our current Government Performance Results Act (GPRA) commitments specify that it is a goal to reduce air toxics emissions by 75% from 1993 levels to significantly reduce American's risk to cancer and other serious adverse health effects caused by airborne toxics. EPA is developing new assessment tools and beginning to address the risk associated with these emissions (as required by the CAA). The EPA will also be modifying that goal that focuses on risk reductions associated with exposure to air toxics. In working toward this risk-based goal, EPA will focus on the cumulative effects of air toxics in urban areas, the multi-media effects of air toxics on water bodies and on populations whose water and food are affected by the deposition of persistent and bio-accumulating air toxics, and the effects on sensitive populations and on economically disadvantaged communities. Eventually, we have a long-term goal of eliminating unacceptable risks of cancer and other significant health problems from exposures to

air toxics emissions and substantially reducing or eliminating adverse effects on our natural environment.

History

In 2000, the EPA introduced the Air Toxics Concept Paper[1], which was a guideline that illustrated EPA's vision on air toxics monitoring in the future. One concept from the Concept Paper was to create a "Pilot" trends program, annual in its duration, which would help the S/L/T and EPA prepare for long term monitoring. The Pilot Program was implemented in 2001, with 10 stations in operation. After the successful implementation of the Pilot Program, Congress granted long term funding for a trends network. Planning for the NATTS began in 2002 with EPA soliciting requests to the States and Local air pollution agencies to work in cooperation in putting together a national trends network. The network has been set at 22 stations throughout the country. The NATTS began sampling January 1, 2004. In 2002-2003, EPA Office of Air Quality and Standards (OAQPS) began to create a Quality System (See Figure 1) that would enhance and allow quality professionals the ability to assess the uncertainty of data collected. A Quality Management Plan, model Quality Assurance Project Plan[2], and a Technical Assistance Guidance[3] document were created to document and disseminate information on the program.

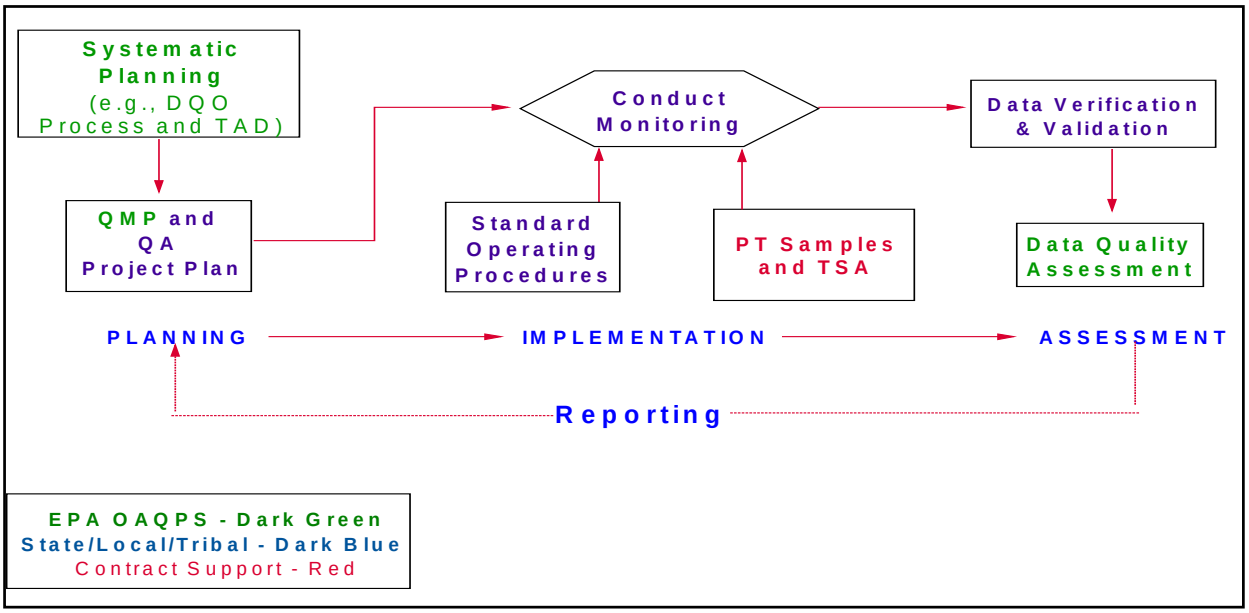


Figure 1. NATTS Quality System

In 2004, OAQPS began planning for the two technical assessment programs illustrated in Figure 1: the Technical Systems Audits (TSAs) and the PT program. The TSAs program began implementation in spring 2004. Battelle Inc., of Columbus Ohio, was selected to perform the TSAs. Approximately one-half of the laboratories and monitoring stations were visited. The results are summarized in the technical paper that is being presented by Battelle Inc. at this conference[4]. The second half of the TSAs will be conducted in the winter and spring 2005.

In fall 2004, the PT program was initiated. Alion Inc., formally known as Mantech Services Inc., was selected to create, distribute and report the findings of the PT program. The first samples for the program were delivered to the laboratories in fall 2004.

The results of the first two sets of data are represented in the following pages.

Results

This section of the paper will discuss the results of the first two sets of PT samples that were reported by Alion Inc. There are three types of samples: Volatile Organic Compounds (VOCs), carbonyls (organic compounds that contain an aldehyde or ketone group) and metals.
Please note the concentrations are "blind", meaning that the laboratory that analyzes the sample will not know the composition or the concentration in each sample. Every laboratory must report their values if they are to participate in the NATTS program. Each of these is prepared in the following manner:

1. VOCs – The contractor notifies the participating laboratories 1 month in advance requesting passivated canisters (6 liter volume). The canister must be cleaned and filled with purified (zero) air. The contractor blends National Institute of Standards and Technology (NIST) traceable compounds from compressed cylinders with zero air in concentrations in the normal analytical range of the program. The canisters are then shipped to the laboratories that analyze the canister in their normal way. The contractor fills two additional canisters during the batch blending and analyzes one canister shortly after all of the canisters are filled and analyzes the second canister after two weeks. The "true" value is the average of the two analyses.
2. Carbonyls - The contractor purchases plastic cartridges that are coated with 2,4 dinitrophenyl hydrazine (DNPH). The cartridges are spiked with carbonyl compounds of interest. Two cartridges are kept for analysis. The cartridges are then shipped to the participating laboratories, which analyze and report their findings.
3. Metals - A subcontractor (Kultech Inc.) is tasked to prepare the metal samples. Kultech Inc. has an instrument that can create metal salt aerosols that can be deposited onto filter media. Kultech creates a NIST traceable solution of metal salts that is nebulized and deposited onto 47-mm filters in batch. The filters are allowed to dry and then are shipped to Alion Inc. Alion analyzes a portion of the PT sample filters and ships the other filters to the participating laboratories.

The following section illustrates the results of the first two PT programs.

Carbonyl Results

Figure 2 illustrates the 4th Quarter formaldehyde results. For this program, the formaldehyde results show consistent results for all laboratories that reported in the third and fourth quarter 2004. Below is a table which illustrates the results of these PTs.

Table 1. Formaldehyde Summary

Test	3rd Quarter 2004	4th Quarter 2004
Mean of % Difference	-1.5%	6.2%
Std Deviation of % Difference.	9.9	17.6

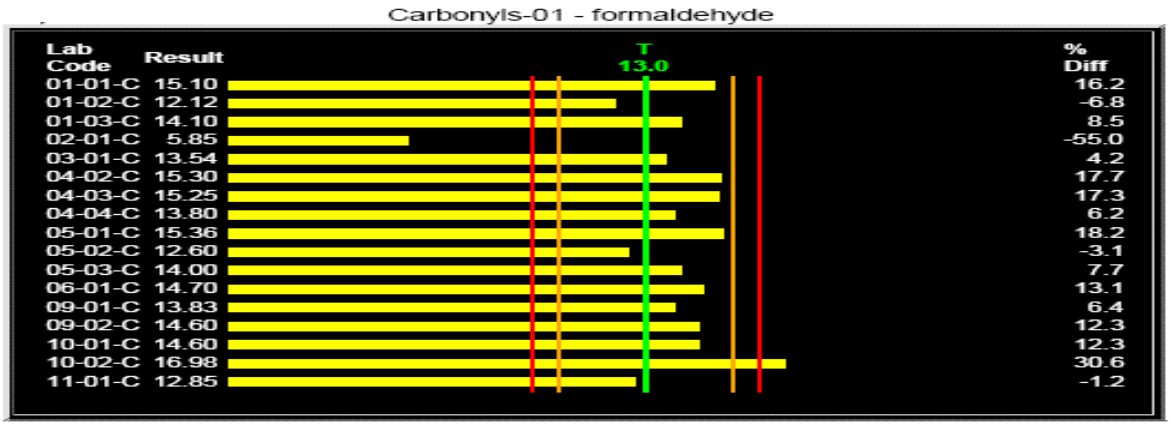


Figure 2. Fourth Quarter 2004 Formaldehyde PT Results

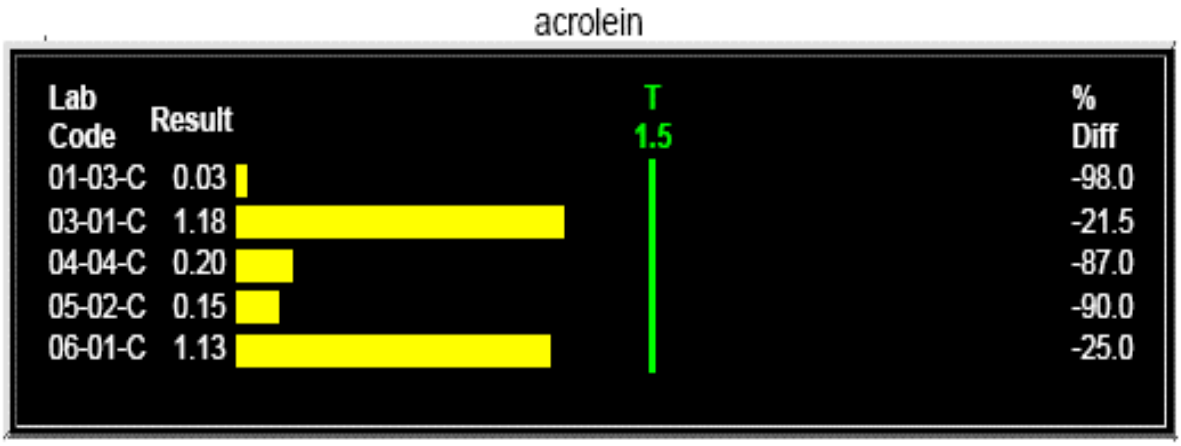


Figure 3. Third Quarter 2004 Acrolein PT Results

Over the two quarters there was only one laboratory that reported values outside of the 25% limit. For acrolein, a compound of interest for the NATTS, recoveries were poor. Of the 16 laboratories that reported, only five reported acrolein. Of these five, only one laboratory reported within the 25% limit. Due to the poor recovery using the DNPH method, acrolein is not being included in the PT program until progress can be made on recovery of this compound.

VOC Results

Figures 4 and 5 illustrate the results of the PT program for the VOCs.

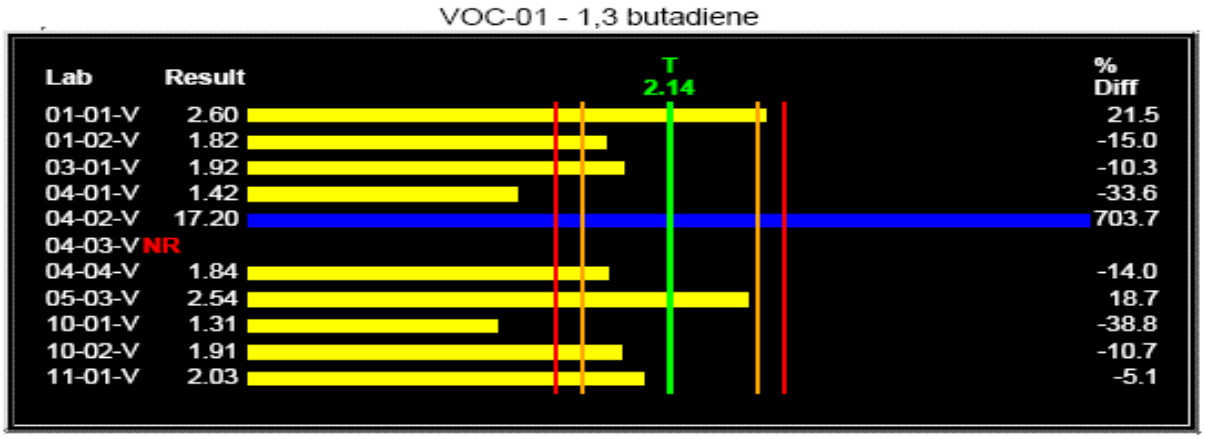


Figure 4. Fourth Quarter 2004 1,3, butadiene PT Results

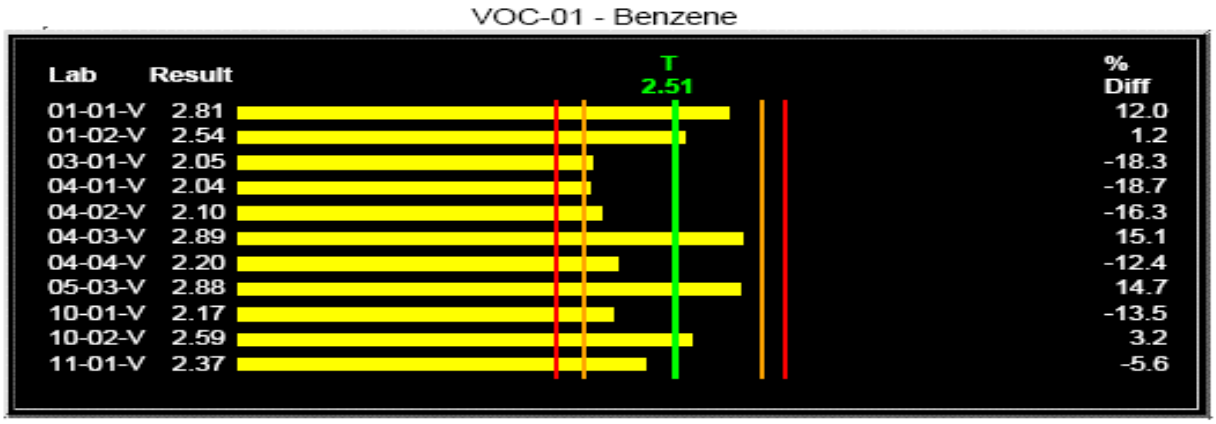


Figure 5. Fourth Quarter 2004 Benzene PT Results

It can be seen from the analysis of these two compounds that there is a discrepancy between reporting. All of the laboratories that participated were within 20% of the calculated value for benzene. However, 1,3-butadiene illustrates a different set of results. One laboratory reported a result that was 703% above the calculated true value and one laboratory did not report the compound at all. This illustrates that there are consistency problems with the laboratories. Other VOC results, although not presented here, illustrate similar results.

Metals Results

The results of the metals PT samples illustrate that there is variability from the reported values from the laboratories. For chromium and lead, the NATTS laboratories that reported were within the 25% limit. However, for cadmium, beryllium and arsenic, not all laboratories were within

the 25% limit. At this time, the deposition procedure used may be causing some of the variability. It will not be known for certain until more PT samples are created and analyzed.

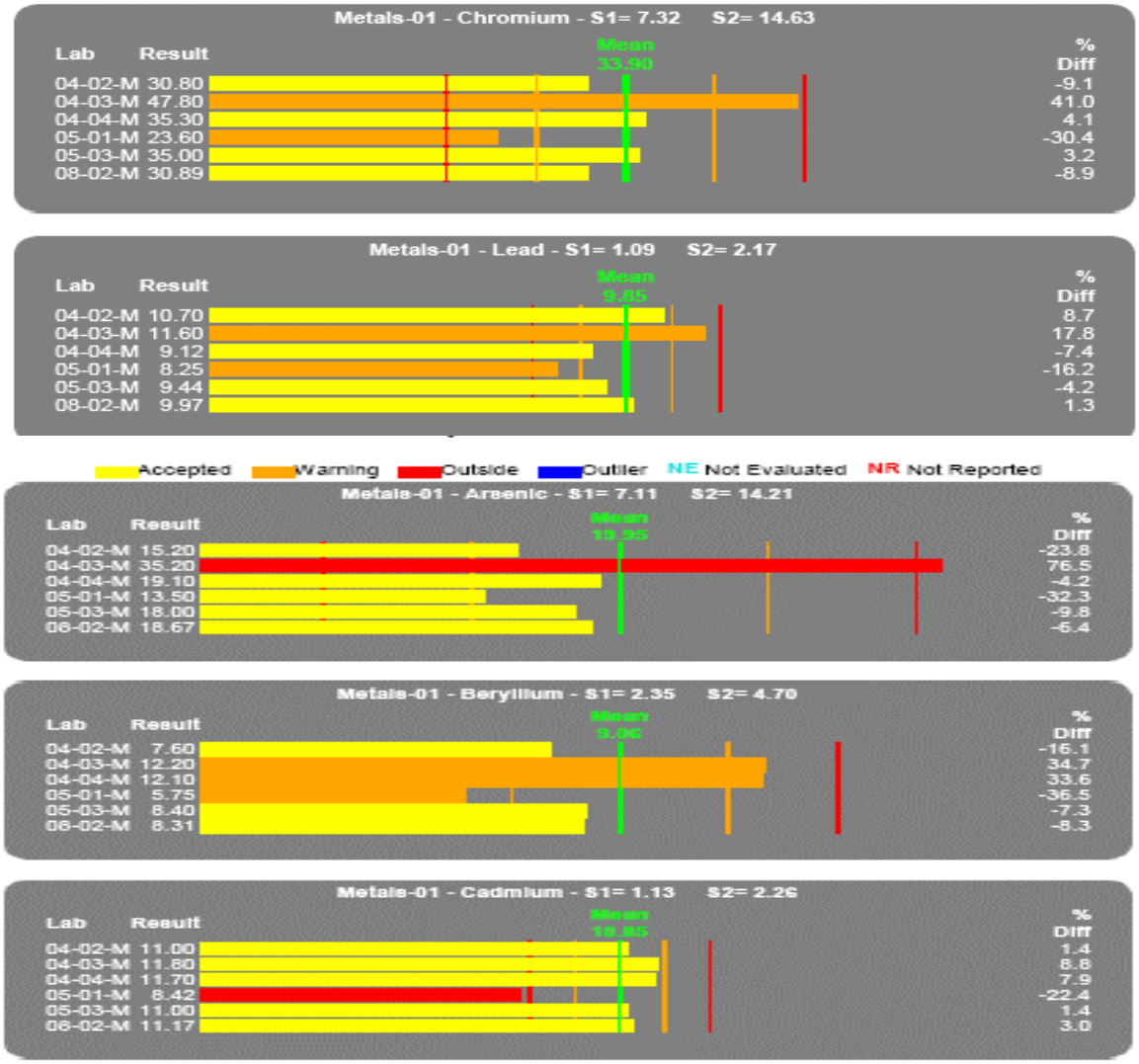


Figure 6. Fourth Quarter 2004 Metals PT Results

Conclusion

The NATTS QA assessment programs began implementation in spring (TSAs) and fall 2004 (PT Samples). This paper focused on the PT program, which provides blind samples to the participating laboratories on a quarterly basis. The PT data will be used to estimate the laboratory bias for the NATTS Program. The results of the third and fourth quarter PTs illustrate that for formaldehyde, the laboratories are able to report values that are within 25% of the calculated true values. For acrolein, a compound of concern and included in the NATTS program, the recovery was poor. Method development for acrolein is on-going and it is hoped that a breakthrough will occur this calendar year. For benzene, cadmium and lead, the reported

values were within the 25% limit. However, this not the case will all VOCs and metals. Some compounds have more variability than others, therefore more work needs to be done to ascertain the causes.

References

1. Air Toxics Monitoring Concept Paper" (Draft), February 29, 2000, http://www.epa.gov/ttn/amtic/airtxfil.html

2. Quality Assurance Guidance Document -- Model Quality Assurance Project Plan for the National Air Toxics Trends Stations, 2002, http://www.epa.gov/ttn/amtic/airtxfil.html

3. National Air Toxics Trends Sites Technical Assistance Document (DRAFT), 2004, http://www.epa.gov/ttn/amtic/airtxfil.html

4. Stetzer,-Biddle, Shannon and Holdren Michael, Technical System Audits (TSAs) and Instrument Performance Audits (IPAs) of the National Air Toxics Trends Sites (NATTS) and Supporting Laboratories, (draft), EPA's 24th Annual National Conference on Managing Environmental Quality Systems - Technical Presentation

Technical System Audits (TSAs) and Instrument Performance Audits (IPAs) of the National Air Toxics Trends Sites (NATTS) and Supporting Laboratories

Shannon Stetzer Biddle and Michael Holdren (Battelle)

Abstract

A National Air Toxics Trends Sites (NATTS) network is being established in the contiguous 48 states by the U.S. EPA's Office of Air Quality Planning and Standards (OAQPS). The NATTS network currently consists of 22 sites (15 urban and 7 rural) throughout EPA Regions I through X. In order to ensure that the data collected are of sufficient quality and to provide a broad overall understanding of the error that is inherent with the data collected for the network, EPA is establishing a Quality System (QS) for the NATTS. Two aspects of the QS are Technical System Audits (TSAs) and Instrument Performance Audits (IPAs) at each network site and at the laboratories that provide the analyses. The U.S. EPA has contracted Battelle to perform all of the audits. Regions I through IV were audited in FY04; Regions V through X are scheduled to be audited in FY05. This paper discusses some of the findings from the audits of the 10 sites and 13 laboratories reviewed in FY04.

Introduction

The current Government Performance Results Act (GPRA) commitments specify a goal of reducing air toxics emissions by 75 percent from 1993 levels to significantly reduce the U.S. population's risk of cancer and other serious adverse health effects caused by airborne toxics. To meet the GPRA goals, a National Air Toxics Trends Sites (NATTS) network is being established in the contiguous 48 states by the U.S. EPA's Office of Air Quality Planning and Standards (OAQPS). The NATTS network currently consists of 22 sites (15 urban and 7 rural) throughout EPA Regions I through X. In order to ensure that the data collected are of sufficient quality to and provide a broad overall understanding of the error that is inherent with the data collected for the network, EPA is establishing a Quality System (QS) for the NATTS. Two aspects of the QS are Technical System Audits (TSAs) and Instrument Performance Audits (IPAs) at each network site and at the laboratories that provide the analyses. The U.S. EPA has contracted Battelle to perform all of the audits. Regions I through IV were audited in FY04; Regions V through X are scheduled to be audited in FY05.

Figure 1 provides a map of all 22 NATTS sites (note that there are actually 23 sites on the map because Tampa, FL, is splitting their money between two monitoring sites). The 10 sites audited in FY04 are displayed in red, the 11 sites scheduled to be audited in FY05 are displayed in blue, and the two monitoring sites that are not being audited by Battelle are displayed in green. (Note that Battelle was not asked to audit the two sites in Region XIII because the Regional offices have already completed similar audits at the sites and there is no need to duplicate the work.) Table 1 explicitly lists each of the sites by audit year and urban or rural classification. This paper discusses the various elements of the audit forms and some of the findings from the 10 sites and 13 laboratories audited in FY04.

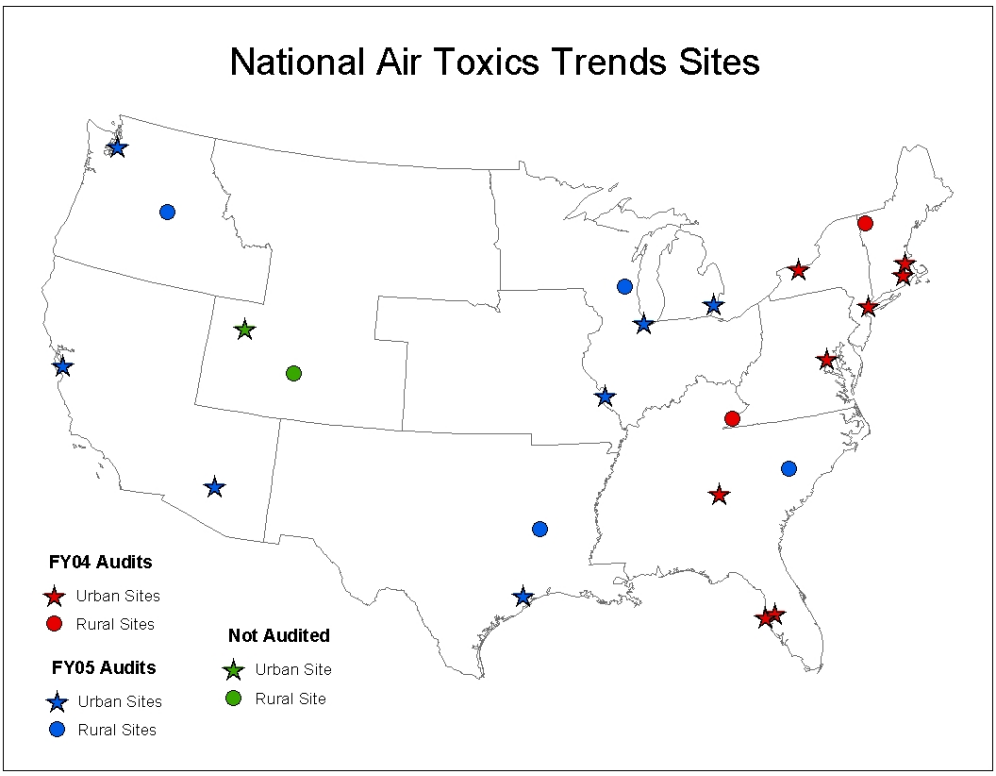


Figure 1. Map of the National Air Toxics Trends Sites.

Table 1. List of NATTS monitoring sites by audit year and urban versus rural.

Audit Year	Urban	Rural
FY04	• E. Providence, RI • Boston (Roxbury), MA • New York (Bronx), NY • Rochester, NY • Washington, DC • Decatur, GA • Tampa, FL (2 sites)	• Chittenden County (Underhill), VT • Hazard, KY
FY05	• Detroit, MI • Chicago, IL • Houston (Deer Park), TX • St. Louis, MO • San Jose, CA • Phoenix, AZ • Seattle, WA	• Chesterfield, SC • Mayville, WI • Harrison County (Karnack), TX • La Grande, OR
Not Audited	• Bountiful, UT	• Grand Junction, CO

TSA and IPA Form Development

Previously generated audit forms provided by OAQPS served as initial guidance for designing the TSA forms used for this program. In addition, the NATTS Technical Assistance Document (http://www.epa.gov/ttn/amtic/files/ambient/airtox/drafttad.pdf) was used directly for generating specific questions regarding the three analytical methods. The TSA is defined as a thorough, systematic, on-site, qualitative audit of facilities, equipment, personnel, training, procedures, record keeping, data validation, chain of custody, data management, and reporting aspects of a system. Separate TSAs were performed at both the monitoring sites and the laboratories. To facilitate auditing, separate TSA forms were developed; one form was used during monitoring site audits and a separate form was used during analytical laboratory audits. The form for the monitoring sites focused on the sampling aspects of the three targeted measurement methods (VOCs, carbonyls, and metals). Similarly, the form for the laboratories covered the analytical aspects of the three methods.

For the field site audits, the IPA was designed to help accomplish two primary objectives. The first of these objectives was to assess sample flow rates through the three sampling systems at the monitoring sites. Calibrated and certified flow rate, temperature, and pressure measuring devices from Battelle's Instrumentation Facility were used in this process. The second objective was to conduct a site assessment at the field monitoring locations. The assessment included such items as making hand drawings of the monitoring site, recording GPS location, taking photos, and logging the distances to obstructions and roadways. CFR 40 Part 58 Appendix E was used for siting assessment guidance.

Tables 2 and 3 provide outlines of the various components of the site TSA/IPA form and the laboratory TSA form, respectively, that were developed by Battelle.

Table 2. Outline of the NATTS field site TSA/IPA form (23 page document).

Part I. General Information Part II. Basic QA/QC A. QAPP and SOPs B. Organization and Responsibilities C. Training, Safety, and Chain-of-Custody D. Sample Handling and Sampling Frequency E. Monitoring Site Housekeeping F. Documentation Part III. Specific Sampling Criteria A. VOC/Canister Sampling B. Carbonyl Sampling C. PM_{10} Metals Sampling	Part IV. Sampler Siting Part V. Instrument Performance Audit A. General B. VOC Sampler C. Carbonyl Sampler D. PM_{10} Metals Sampler

Table 3. Outline of the NATTS analytical laboratory TSA form (37 page document).

Part I. General Information Part II. Basic QA/QC A. QAPP and SOPs B. Organization and Responsibilities C. Quality Assurance/Quality Control D. Training E. Safety F. Document Control and Records G. Facilities, Equipment, Software Part III. VOC/Canister Analysis A. Canister Cleaning Equipment B. Canister Cleanliness C. Canister Analysis Procedures D. Chain-of-Custody and Sample Handling E. Performance Evaluation	Part IV. Carbonyl Analysis A. Carbonyl Sampler Cleanliness B. Analysis Procedures C. Chain-of-Custody and Sample Handling D. Performance Evaluation Part V. PM_{10} Metals Analysis A. Filter Preparation B. Sample Receipt & Storage C. Sample Digestion D. Metals Analysis E. Chain-of-Custody and Sample Handling F. Performance Evaluation

FY04 Audit Summaries

Overall, the NATTS project managers and supporting technical staff members audited in FY04 were very conscientious and well qualified, as were the staff from the supporting analytical laboratories. The audit team members found very few issues of concern. Significant issues were called out in the individual audit reports and in almost all cases, corrective action was taken immediately.

One primary result of the TSAs was the revelation of a number of commonalities and weaknesses among sites and laboratories. For example, while a majority of the sites and laboratories audited in FY04 had updated Quality Assurance Project Plans (QAPPs) and Standard Operating Procedures (SOPs) in place, a few field sites and laboratories were operating without these critical documents in place. Table 4 provides a summary of a few important findings. One finding in particular revealed that while all ten field monitoring sites had QAPPs present and up to date, only 11 of the 13 analytical labs had QAPPS present and up to date. Similar deficiencies were observed with SOPs being present and up to date. Furthermore, only six of the 13 analytical labs had a formal controlled document program in place. In fact, at least one location was found to be using "old" SOPs. Findings also indicated that internal audits were being performed at only four of the 10 monitoring sites and at seven of the 13 analytical labs.

Table 4. Examples of commonalities and weaknesses among sites and laboratories.

	The number of sites/laboratories where…			
	…the QAPP is present and up to date.	**…the SOP(s) is/are present and up to date.**	**…a formal controlled document program is in place.**	**…internal audits are performed.**
Monitoring Sites (10 total)	10	9	NA[1]	4
Analytical Labs (13 total)	11	11	6	7

[1] Did not ask this question at the monitoring sites.

Highlights of Site and Laboratory Assessments

In this section, information is presented showing the environmental differences between the urban and rural site locations, the types of samplers used for each measurement method and the results from the instrument flow checks. A summary of the analytical laboratories associated with the various sites is also shown.

Monitoring site photos are shown in Figures 2 and 3 to convey the environmental differences between the audited urban and rural monitoring locations. Figure 2 provides pictures taken at the Roxbury, Massachusetts, urban field site; this location is an environmental justice site. The monitoring equipment is contained within a chain-link fenced area, bordered to the north by an automobile salvage yard, to the west by a series of stores, to the south by an electrical utility station, and to the east by a street. In contrast, Figure 3 shows pictures taken at the Hazard, Kentucky, rural field site. Here the monitoring equipment is contained within a chain-link fenced area on the top of a large hill, with a school located below to the north/northwest of the site. The area around the trailer and fence is clear for approximately 100 feet in all directions and the land outside of that area is covered in trees.

Figure 2. Example of an urban monitoring location – Roxbury, Massachusetts – looking to the East and looking to the South, respectively.

Figure 3. Example of a rural monitoring location – Hazard, Kentucky – looking to the West and looking to the North, respectively.

Within each of the sites, three monitoring methods were audited as part of the quality system program: VOCs, carbonyls, and PM_{10} metals. Across sites, more than one type of sampler manufacturer was operated for each method. Table 5 summarizes the different manufacturers and the number of corresponding sites. The table shows that a variety of canister sampler manufacturers are used, whereas only two manufacturers exist for carbonyl and PM_{10} samplers. [Note that while 10 sites were audited, only nine sites were collecting PM_{10} metals, thus, only nine samplers are accounted for in the metals row of Table 4.]

Table 5. Summary of sampler manufacturers by sampling method.

Sampler Method	Sampler Manufacturer (n = number of sites out of 10)[a]
Canister	• Xontech (n=4) • ATEC (n=2) • In-house design (n=2) • Andersen AVOCs (n=1) • Meriter (n=1)
Carbonyl	• ATEC (n=8) • ERG (n=2)
PM_{10} Metals	• Andersen type (n=7) • Wedding (n=2)

[a] One monitoring site had not begun sampling metals at the time of the audit, thus, the total number of sites monitoring PM_{10} metals is nine.

The technical systems audits also revealed that a number of sites were collecting either collocated or duplicate samples with at least one sampling method. Table 6 shows the number of monitoring sites collecting collocated and duplicate samples. For example, one site is collecting collocated canister samples (i.e., two samplers), another site is collecting duplicate canister samples (i.e., one sampler, two channels), and the remaining eight sites are collecting a single sample. This means that while a total of 10 sites were being audited, there were a total of 11 canister samplers, or 12 channels, in operation. Similarly, there were 11 carbonyl samplers (16 channels), and 12 PM_{10} samplers operating across the 10 sites.

Table 6. Number of monitoring sites collecting collocated or duplicate samples.

Sampler Method	Number of Sites (out of 10 total) [a]		
	Collocated	Duplicate	Single
Canister	1	1	8
Carbonyl	1	5	4
PM_{10} Metals	3	NA	6

[a] One monitoring site had not begun sampling metals at the time of the audit, thus, the total number of sites monitoring PM_{10} metals is nine.

As an integral part of the instrument performance audit conducted at each monitoring site, flow checks were performed on all three sampling methods. For the 10 monitoring sites audited in FY04, the flow rate comparisons look reasonably good, especially for the carbonyl sampling method. As shown in Table 7 and Figure 4, all differences in the carbonyl flows are less than 10 percent. Most canister sampling values are within five percent, with two apparent outliers above 10 percent. Finally, the PM_{10} sampler values are also primarily less than ten percent, however, there are three values greater than 10 percent and one of those values is a large outlier at approximately 19 percent. [Note that "n" equals the number of flow audits included in the summary.] In summary, only 5 of the 37 flow rate checks were above 10 percent.

Table 7. Summary of the FY04 NATTS IPA flow results by sampling method.

Percent Difference	Canister Sampler (n=9)[a]	Carbonyl Sampler (n=16)	PM_{10} Sampler (n=12)
Min	0.1	0.2	0.1
Median	2.7	5.6	4.25
Max	14.0	9.7	19.3

[a] Unable to audit the canister sampling units at the Tampa, Florida, sites because the in-house designed field flow controllers were not compatible with audit flow devices.

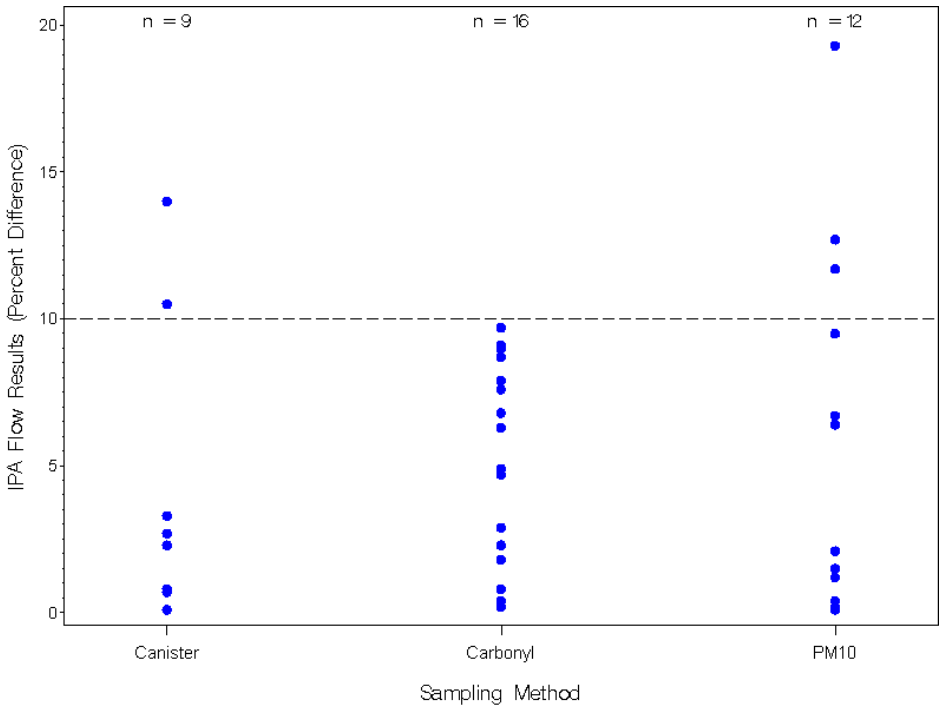


Figure 4. Distribution of the FY04 NATTS IPA flow results by sampling method.

Thirteen (13) different laboratories were included in the FY04 audits; this means that there was not always one laboratory assigned to each site. In fact, some sites had one laboratory perform all three analyses (e.g., Decatur, Georgia), some sites had two analyses done by one laboratory and one analysis done by another laboratory (e.g., Providence, Rhode Island), and other sites had a separate laboratory for each sampling method (e.g., Tampa, Florida). A complete listing of the analytical laboratories responsible for analyzing the samples collected at the respective NATTS monitoring sites is presented in Table 8.

Table 8. Summary of the laboratories responsible for analysis of the samples collected at NATTS monitoring sites audited in FY04.

EPA Region	NATTS Monitoring Site	Monitoring Method[1]	Analytical Laboratory
Region I	Roxbury, MA	V	RI Dept. of Health
		C	MA Dept. Envir. Protection (in PAMS season)
			RI Dept. of Health (in off-season)
		M	ERG/RTI
	Chittenden, VT	V, C, M	VT Dept. of Envir. Conservation[2]
	Providence, RI	V, C	RI Dept. of Health
		M	USEPA Region 1 - New England Regional Laboratory (temporary arrangement)[3]
Region II	Rochester, NY and Bronx, NY	V, C	NY Dept. of Envir. Conservation
		M	NYS Dept. of Health [3, 4]
Region III	Washington, DC	V	MD Dept. of the Environment
		C	Philadelphia Dept. of Public Health
		M	WV Dept. of Envir. Protection
Region IV	Decatur, GA	V, C, M	GA Dept. of Natural Resources Envir. Protection Division
	Hazard, KY	V, C, M	KY Division of Envir. Services
	Tampa, FL: Pinellas County and Hillsborough County	V	Pinellas County Dept. of Envir. Management
		C	ERG
		M	Envir. Protection Commission of Hillsborough County

[1] V=VOCs/Canisters, C=Carbonyls, M=Metals (PM_{10}).
[2] The VT DEC was doing VOC and carbonyl analyses at the time of the audit, and planned to implement ICP-MS metals analyses.
[3] Did not audit.
[4] At the time of the New York audits, there was some uncertainty about which laboratory would be awarded the contract to do the metals analyses.

Analytical methods associated with the canister, carbonyl, and PM_{10} metals samplers were audited at the supporting laboratories. One thing the TSAs confirmed was that the TO-15 method and the TO-11 method were exclusively followed for canister and carbonyl analysis, respectively. The TSAs also revealed that the metals analyses were conducted using inductively coupled plasma-mass spectrometry (ICP-MS) at four of the five audited metals laboratories; however, the remaining metals laboratory used inductively coupled plasma-optical emission spectroscopy (ICP-OES). This procedure is not the recommended procedure in the NATTS technical assistance document.

Summary

The first year's audits of the NATTS and supporting laboratories focused on U.S. EPA Regions I through IV. A total of twenty-three final audit reports were generated. Overall, the NATTS project managers and supporting technical staff members were very conscientious and well qualified to conduct air monitoring efforts using the canister, carbonyl, and PM_{10} sampling devices. The staff from the analytical laboratories likewise exhibited considerable expertise during the audit sessions. The audit team members found very few issues of concern. Significant issues were called out in the individual audit reports and in almost all cases, corrective action was taken immediately.

Interlaboratory Comparison of Ambient Air Samples

Charles L Pearson

California Air Resources Board
Quality Assurance Section
1927 13th Street
Sacramento, CA 95814

ABSTRACT

The California Air Resources Board (CARB) has developed a specialized sampling technique to collect ambient air samples that are used to evaluate the relative differences in non-methane hydrocarbon (NMHC) species among Photochemical Assessment Monitoring Station (PAMS) laboratories and toxics species amid toxics laboratories. The sampling system is engineered to simultaneously collect several ambient air samples from a site with historically high concentrations of hydrocarbon or toxics species. The specialized sampling unit is capable of filling up to 14 canisters simultaneously, with a canister being sent to each of the participating designated PAMS or toxics laboratories for analysis. Each laboratory follows their standard operating procedure in assaying the contents from the comparison check canister and reports a value for each detected compound to CARB. The laboratory responses are then tabulated and rigorous statistical tests are performed on each of the values to achieve an accurate depiction of the canisters' contents. The interlaboratory comparison check allows us to assess the variability of the measurement process using real-world samples at ambient level concentrations. This comparison check method emulates a round-robin check using a single canister, however, it does not encounter delays in canister routing between laboratory participants or experience gradual loss of pressure as the canister contents are analyzed.

The following paper explains the equipment, sampling methodology, and statistical techniques developed for the ambient air interlaboratory comparison check. This paper also details the history, as well as future applications of the interlaboratory comparison procedure.

INTRODUCTION

To evaluate the accuracy of data generated by the PAMS and toxics laboratories, CARB conducts annual interlaboratory comparison checks. The interlaboratory comparison check is one of many quality assurance tools used to assess data quality and evaluate laboratory practices. The comparison check complements the laboratory and through-the-probe audit programs by evaluating the performance of the participating laboratories relative to one another using a real-world air sample. The purpose of the laboratory comparison is to indicate general agreement or not among the laboratories, and is not

necessarily an indication of accuracy. The comparison check program was initiated in 1998 with its focus on NMHC species. A paper titled "Interlaboratory Comparison of Ambient Air Samples" was presented at the United States Environmental Protection Agency's Air and Waste Management Association Symposium in September 2000.[1] Since 2000, CARB expanded the program to support the toxics program, updated the sampling equipment to sustain additional laboratories, and enhanced the statistical analysis of the data.

The comparison is intended to support each program by evaluating the ability of each laboratory to produce consistent data from an ambient air sample in terms of the number of compounds present and their concentrations. It also enables laboratories to directly compare their responses with other laboratories located throughout the United States. The interlaboratory comparison protocol is similar to a round-robin check, with the primary difference being that each participating laboratory receives a separate canister. This comparison check procedure is more effective than a round-robin check as it does not encounter delays in canister routing between laboratory participants or experience gradual loss of pressure as the canister contents are analyzed. Since a multiple canister approach is used, the samples experience limited travel time between collection and analysis. The sampling time and location are based on historically high temporal and spatial concentrations of pollutants. Typically, the ideal sampling location and time has been a site with close proximity to freeways during early morning hours to capture commute patterns.

Once the samples are collected, each laboratory conducts a minimum of two analyses from the canister contents using a gas chromatograph (GC) to determine the compounds present and their concentrations. Each laboratory reports its results to CARB. CARB in turn tabulates the responses and calculates the mean and standard deviation for each compound. Using the mean and standard deviation, upper and lower critical values are established to identify outliers in the data set. Responses that exceed either critical value are eliminated and an adjusted mean and standard deviation are then calculated. Each individual laboratory response is compared against the adjusted mean and standard deviation. The laboratories are notified of any response that differs more than two standard deviations from the adjusted mean response.

SAMPLING PROCEDURE

Background

Each participating laboratory is required to provide one clean, evacuated, 6-liter stainless steel canister. One laboratory is asked to provide two canisters and they are requested to analyze both cans to verify the precision of the collection procedure. The sample canisters from all laboratories are simultaneously filled using a modified canister sampler with ambient air over a three-hour period. The three-hour sampling period ensures that a representative sample is collected. Following sample collection, the canisters are

returned to their respective laboratories for GC analysis. Results from the analyses are then forwarded to CARB, which compiles the results and performs an assortment of statistical calculations on each of the reported values. The compiled results, in tabular and graphical form, are distributed to each participating laboratory.

Equipment

A specialized RM Environmental Systems Inc. (RMESI) 910A™ canister sampler is equipped with a larger stainless steel pump, with Viton rings, as well as, a 2000 cubic centimeters per minute (cc/min) mass flow controller (MFC) to supply the increased flow required when filling numerous canisters. A custom built manifold was constructed to accept up to fourteen, 6-liter sampling canisters to be filled from the sampler's single inlet port. The collection probe inlet consists of ¼ inch stainless steel tubing, with ⅛ inch stainless lines supplying the sample to each of the canisters. Certified temperature, pressure and RH sensors are used to record ambient meteorological data.

Pre-Sample Cleaning

All sampling equipment, including the probe inlet and external presentation lines, are cleaned prior to sampling by flushing the entire system using zero air and ultrapure nitrogen. Zero air is passed through the system for a minimum of eight hours, followed by an ultrapure nitrogen purge for three hours. Once the three-hour nitrogen purge is complete, a certified clean, evacuated canister is connected to one of the sampler's output lines. The sampler is then allowed to draw in ultrapure nitrogen until the canister reaches a pressure between 12-14 psig. The sample canister is analyzed to insure that no contamination exists in the sampling system. If the analysis results indicate contamination, the system purge must be repeated. Once the sampling system cleanliness has been certified, all sampler ports and lines are capped to maintain an uncontaminated system. Each participating laboratory is required to submit one or two certified clean canisters following their normal canister cleaning procedure. CARB also requests documentation of the cleanliness of their canister. Since canister cleanliness represents a variable in the comparison check procedure, the comparison also serves as an indirect check of the canister cleaning process.

System Set-up

Once the sample site is selected, the sample probe inlet is situated in the same position as the probe routinely used at the station for sampling. The inlet probe to the sampler is connected to a ¼ inch stainless steel probe line. The sample canisters are then connected to the sampler outlet ports using ⅛ inch stainless steel tubing. Figure 1 illustrates a schematic of the system set-up. Prior to sample collection, a leak-free sampling system must be achieved. To perform the required leak check, the valve on one canister is opened, causing the gauge on the sampler to register a vacuum approximately equal to that of the opened canister, the valve is then immediately closed. If the system has maintained the initial vacuum after 15 minutes, the system is considered to be leak-free. If the system does not hold vacuum for the 15 minute period, the canister fittings and their associated connections are re-tightened and the leak check is performed again. The sampling unit's flow rate setting is determined by using the following equation. Using

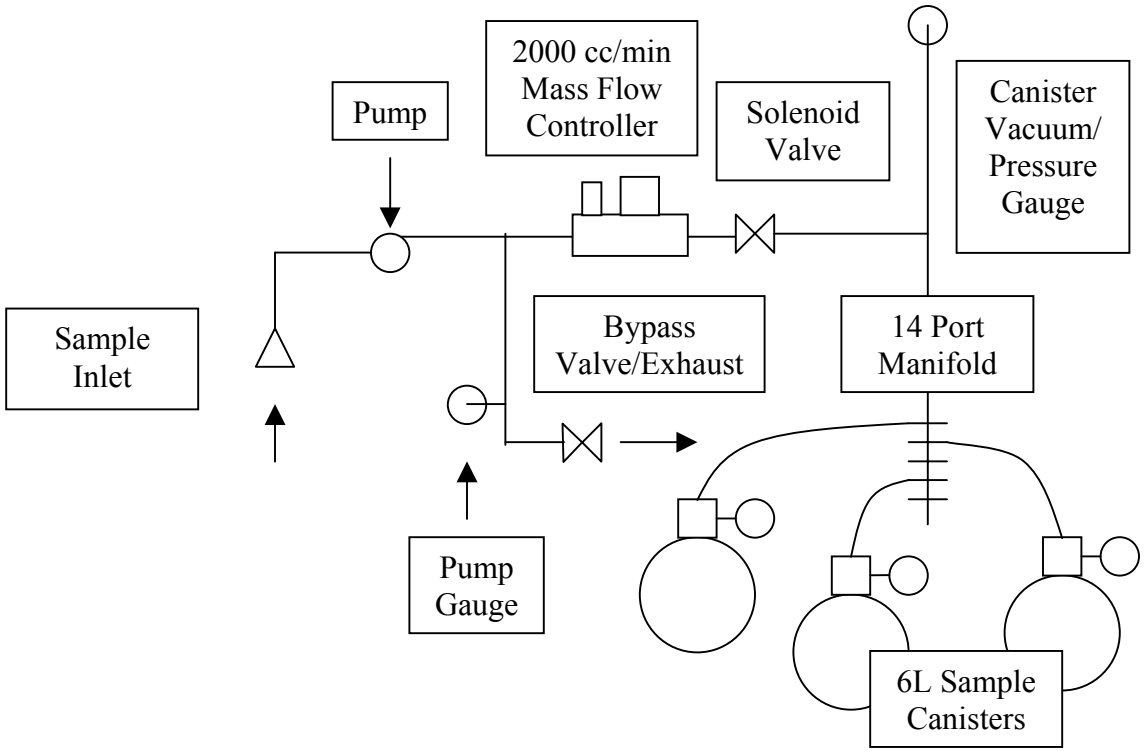
Equation 1, the sampler's flow rate for filling 14 canisters in 3 hours, each with a final pressure of 2 atmospheres, should be adjusted to 933 cc/min.[2]

$$F = N [(P) (V)] / (T) \qquad \text{[Equation 1]}$$

$$F = 14 [(2 \text{ atm}) \times (6000 \text{ cc/atm})] / (180 \text{ min})$$

Where: F = flow rate, in cubic centimeters per minute (cc/min)
N = number of canisters
P = final canister pressure, in atmospheres absolute (atm)
V = volume of sample canister, in cubic centimeters (cc)
T = desired sampling time, in minutes (min)

Figure 1. Sampling System Set-up



Sample Collection

Once the sampling unit is energized and prior to the start of sampling, the unit has a required 30-minute internal air purge. During this time all canister valves are opened and each canisters pre-sample vacuum is recorded. Once the internal air purge is complete, a solenoid actuates which allows airflow to run through the entire manifold and to the canisters. When the canister pressures reach approximately 14 psig, the canister valves are closed and the sampler power turned off. The canisters are then removed from the sample lines, capped, and stored properly. The filled canisters are then returned to their respective laboratories for analysis.

Results

Following their established standard operating procedures, each laboratory conducts a minimum of two analyses from their canister and reports its results, which includes the average for each detected species. Additionally, each laboratory must indicate the limit of detection of their instrument. After CARB has received each laboratory's results, the data are reviewed to detect any probable anomalies. Any response that appears to be a possible abnormality is flagged. All flagged data are confirmed with the reporting laboratory. The data are then tabulated and the average concentration and standard deviation for each compound are calculated. To eliminate atypical, infrequent values (outliers) from being included in the statistical calculations, upper and lower critical values are established. The critical values are based upon the probability that 80% of these values will fall within this range (two-tailed test). [3] All laboratory responses that exceed the upper or lower critical value are not included in the adjusted mean or standard deviation calculation. Prior to using critical values to determine outliers, the data analysis used a standard range for determining an adjusted mean, which created bias in responses with low concentrations. Since critical values are calculated directly from each compound's average and standard deviation, the adjusted mean and standard deviation are more accurately represented. By applying Equation 2, the critical values for each compound are calculated and established.

Upper Critical Value = St Dev x 1.28 + Mean [Equation 2]
Lower Critical Value = St Dev x 1.28 - Mean

Where: St Dev = standard deviation of all responses for each compound
Mean = mean of all responses for each compound, in ppb

Each laboratory response is then compared against the adjusted mean response. The results are compiled in a table that includes the mean, standard deviation, adjusted mean, adjusted standard deviation, and critical values for each target compound (Table 1). Graphs are generated depicting each laboratory response for each compound with the adjusted mean response for all laboratories, as well as, graphs displaying all laboratory responses (Figures 2 and 3). Prior to 2003, the graph plot error bars indicated a range of +/-20% from the adjusted mean response to illustrate laboratory performance. This approach created bias in compounds with low mean concentrations. To achieve a more accurate portrayal of laboratory performance, CARB implemented error bars using +/-2 standard deviations from the adjusted mean for each compound. The table and graph plots allow each laboratory to compare its responses to the responses from all other participants. The purpose of the laboratory comparison is to indicate general agreement or not among the laboratories, and is not necessarily an indication of accuracy. Each laboratory receives a result letter that explains the interlaboratory comparison process and details all reported responses that differ more than two standard deviations from the adjusted mean response for any given compound.

CONCLUSION

The ambient air interlaboratory comparison check program has steadily progressed since its inception. CARB has improved the program by updating and utilizing superior equipment, expanding participation, and making the results more useful for its participants. In 1998, the first comparison check was conducted with a few California laboratories using a sampler built by CARB staff with compression fittings and stainless steel tubing. The early comparisons had a small sample size and included limited data and statistical analysis. Implementing a new sampler with greater capability has allowed CARB to vastly increase the number of laboratory participants. Presently, the comparison has expanded to include up to 14 laboratories with locations nationwide. The use of greater statistical methods improves the results and allows each laboratory to see how they compare with others when assaying ambient air. CARB is continuously striving to improve the comparison with future improvements and developments. Some of the future improvements may include increasing the sampler flow capability and sampling ports potential, initializing interlaboratory comparison checks within other programs, and investigating the possibility of introducing a NIST certified tracer compound into the sampling stream.

REFERENCES

1. Webster S. Tasat and Eric V. Albright; "Interlaboratory Comparison of Ambient Air Samples", Presented at the U.S. EPA/Air & Waste Management Association International Symposium on the Measurement of Toxic and Related Air Pollutants, Research Triangle Park, North Carolina, September 12-14, 2000.

2. "Compendium of Methods for the Determination of Toxic Organic Compounds in Ambient Air: Method TO-15, Second Edition", U.S. Environmental Protection Agency, Cincinnati, OH, EPA/625/R-96/010b, January 1997.

3. Trudy A. Watt, "Introductory Statistics for Biology Students 2nd Edition", Chapmann and Hall/CRC, Boca Raton, Florida, 1997

Table 1.

2004 Ambient Air Toxics Laboratory Comparison Check

Compound	Lab 1 (ppbv)	Lab 2 (ppbv)	Lab 3a (ppbv)	Lab 3b (ppbv)	Lab 4 (ppbv)	Lab 5 (ppbv)	Lab 6 (ppbv)	Lab 7 (ppbv)	Lab 8 (ppbv)	Lab 9 (ppbv)	Lab 10 (ppbv)	Lab 11 (ppbv)	Lab 12a (ppbv)	Lab 12b (ppbv)	Lab 12c (ppbv)	Lab 12d (ppbv)	Lab 12 Avg (ppbv)	Total	Mean	St Dev	Lower Critical Value	Upper Critical Value	Adj Mean	St Dev of Adj Mean	St Dev from Adj Mean	
																									-2 SD	+2 SD
1,4-dichlorobenzene	--	0.07	--	--	0.03	0.04	0.04	0.03	--	--	--	--	--	--	--	--	--	0.21	0.04	0.02	0.02	0.06	0.04	0.01	0.02	0.06
1,1,1-trichloroethane	--	0.04	--	--	0.03	--	0.02	0.02	--	--	--	0.03	0.03	0.03	0.03	0.03	0.03	0.14	0.03	0.01	0.02	0.04	0.03	0.01	0.01	0.05
1,3-butadiene	--	--	--	--	0.02	--	0.02	0.04	--	--	0.03	0.03	--	--	--	--	--	0.11	0.03	0.01	0.02	0.04	0.03	0.01	0.01	0.05
carbon tetrachloride	--	0.11	0.08	0.08	0.08	0.16	0.08	0.07	--	--	0.10	0.10	0.13	0.14	0.13	0.14	0.14	1.00	0.10	0.03	0.06	0.14	0.09	0.02	0.05	0.13
chloroform	--	0.04	--	--	0.03	0.06	0.03	0.03	--	--	0.06	0.02	0.04	0.04	0.04	0.03	0.04	0.30	0.04	0.01	0.02	0.06	0.04	0.01	0.02	0.06
1,3,5-trimethylbenzene	--	0.03	--	--	0.08	0.13	0.07	0.05	--	--	--	--	--	--	--	--	--	0.36	0.07	0.04	0.02	0.12	0.06	0.02	0.02	0.10
trichlorotrifluoroethane	--	--	--	--	0.08	0.15	--	--	--	--	--	0.08	--	--	--	--	--	0.31	0.10	0.04	0.05	0.16	0.10	0.04	0.02	0.18
trichloroethylene	--	0.03	--	--	0.02	--	0.02	0.02	--	--	0.01	0.02	0.03	0.03	0.03	0.03	0.03	0.13	0.02	0.01	0.01	0.03	0.02	0.01	0.00	0.04
styrene	--	0.03	--	--	0.04	0.07	0.04	0.03	--	--	0.05	--	--	--	--	--	--	0.26	0.04	0.02	0.02	0.06	0.04	0.01	0.02	0.06
tetrachloroethylene	--	0.09	--	--	0.08	0.13	0.09	0.07	--	--	0.05	0.09	--	--	--	--	--	0.60	0.09	0.02	0.05	0.12	0.08	0.02	0.04	0.12
bromomethane	--	0.04	--	--	0.02	--	--	0.03	--	--	--	--	--	--	--	--	--	0.09	0.03	0.01	0.02	0.04	0.03	0.01	0.01	0.05
dichlorodifluoromethane	0.60	0.62	--	--	0.56	1.13	--	--	0.54	0.60	0.64	--	--	--	--	--	--	4.69	0.67	0.21	0.41	0.93	0.59	0.04	0.51	0.67
o-xylene	--	0.16	0.13	0.19	0.20	0.38	0.22	0.13	0.16	0.16	0.20	0.20	0.21	0.19	0.22	0.20	0.21	2.34	0.19	0.07	0.11	0.28	0.18	0.03	0.12	0.24
chloromethane	0.60	0.49	--	--	0.53	0.93	0.58	0.49	0.55	0.56	--	--	--	--	--	--	--	4.73	0.59	0.14	0.41	0.77	0.54	0.04	0.46	0.62
benzene	--	0.40	0.38	0.32	0.34	0.74	0.37	0.30	0.29	0.36	0.37	0.40	0.33	0.37	0.34	0.38	0.36	4.63	0.39	0.12	0.24	0.54	0.35	0.04	0.27	0.43
dichloromethane	0.60	0.66	0.52	0.54	0.59	0.98	0.59	--	0.73	0.63	0.91	0.70	0.64	0.61	0.63	0.59	0.62	8.07	0.67	0.14	0.49	0.85	0.62	0.07	0.48	0.76
trichlorofluoromethane	--	0.30	--	--	0.29	0.55	--	--	0.30	0.31	0.23	0.26	--	--	--	--	--	2.24	0.32	0.11	0.19	0.45	0.28	0.03	0.22	0.34
ethylbenzene	--	0.18	0.17	0.13	0.16	0.32	0.21	0.12	0.15	0.16	0.22	0.15	--	--	0.20	--	0.20	2.17	0.18	0.05	0.11	0.25	0.17	0.03	0.11	0.23
m/p-xylene	--	0.30	0.65	0.48	0.47	0.75	0.55	0.34	0.34	0.35	0.67	0.40	0.50	0.47	0.57	0.51	0.51	5.82	0.48	0.15	0.30	0.67	0.46	0.13	0.20	0.72
propylene	--	--	--	--	--	0.75	0.42	--	--	--	--	--	--	--	--	--	--	1.17	0.59	0.23	0.29	0.88	0.59	0.23	0.13	1.05
1,2,4-trimethylbenzene	--	0.09	--	--	0.32	0.41	0.29	0.23	0.22	0.16	0.29	--	--	--	--	--	--	2.01	0.25	0.10	0.12	0.38	0.23	0.08	0.07	0.39
methyl ethyl ketone	--	1.33	1.25	0.98	--	2.01	1.05	--	0.87	--	--	0.80	--	--	--	--	--	8.29	1.18	0.41	0.66	1.71	1.05	0.21	0.63	1.47
toluene	2.80	1.82	2.41	2.25	2.50	4.24	2.40	2.61	2.40	2.34	3.20	1.90	2.46	2.41	2.47	2.50	2.46	33.33	2.56	0.61	1.78	3.35	2.42	0.36	1.70	3.14

Note: (1) -- = Not reported.

(2) = Not included in statistical calculations at laboratory's request

(3) = Responses exceeded the upper or lower critical value and were not included in the adjusted mean or standard deviation.

(4) The lower and upper critical values were established to exclude outliers when comparing all laboratory responses.

For each compound:

Lower critical value = St Dev X 1.28 - mean

Upper critical value = St Dev X 1.28 + mean

(5) Labs 1, 2, 6 and 12 reported values were adjusted to two significant figures.

(6) Lab 3a and 3b were two separate canisters analyzed by two different instruments.

(7) Lab 12 analyzed collocated canisters with two different instruments, the average value was used for statistical calculations.

Figure 2.

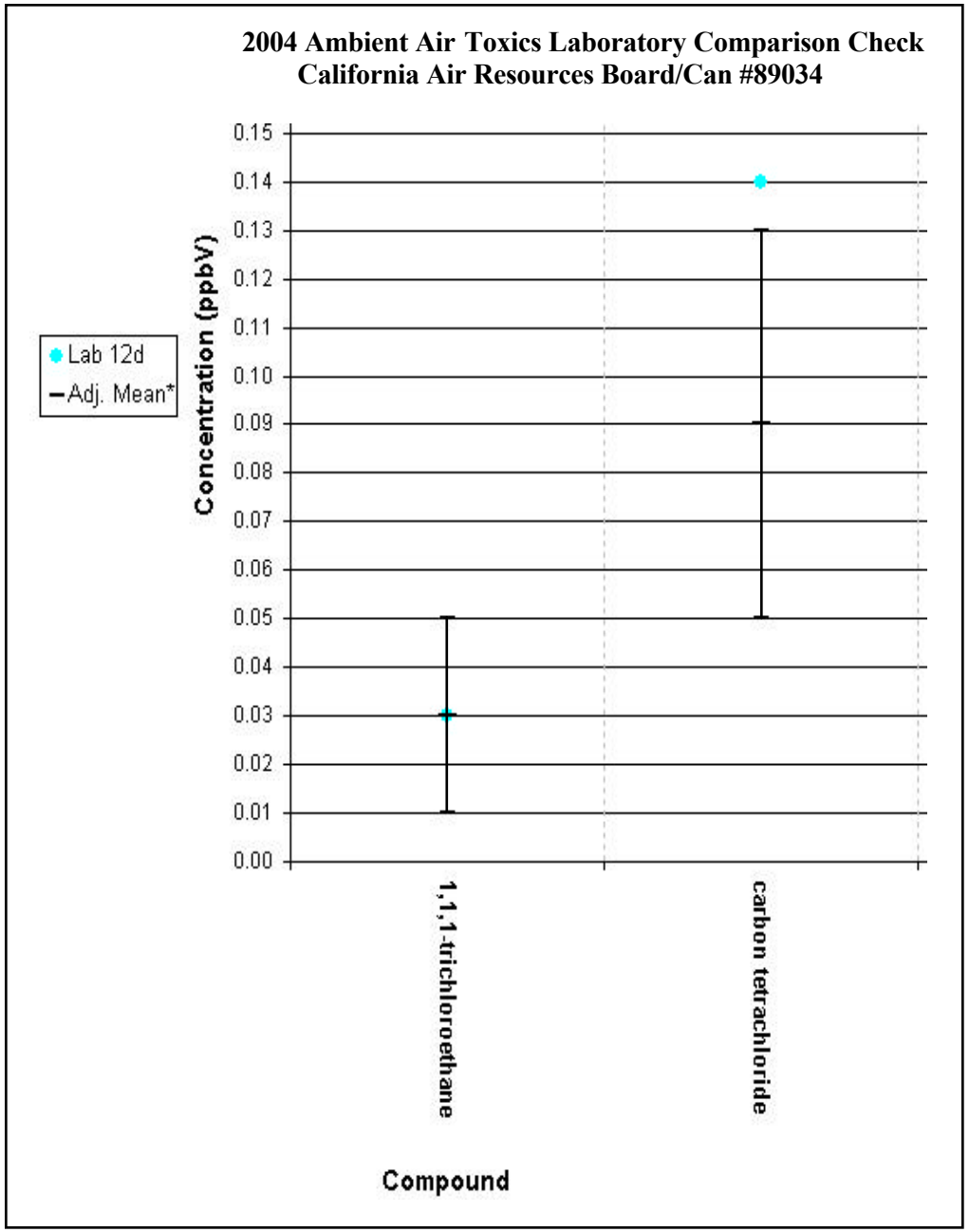
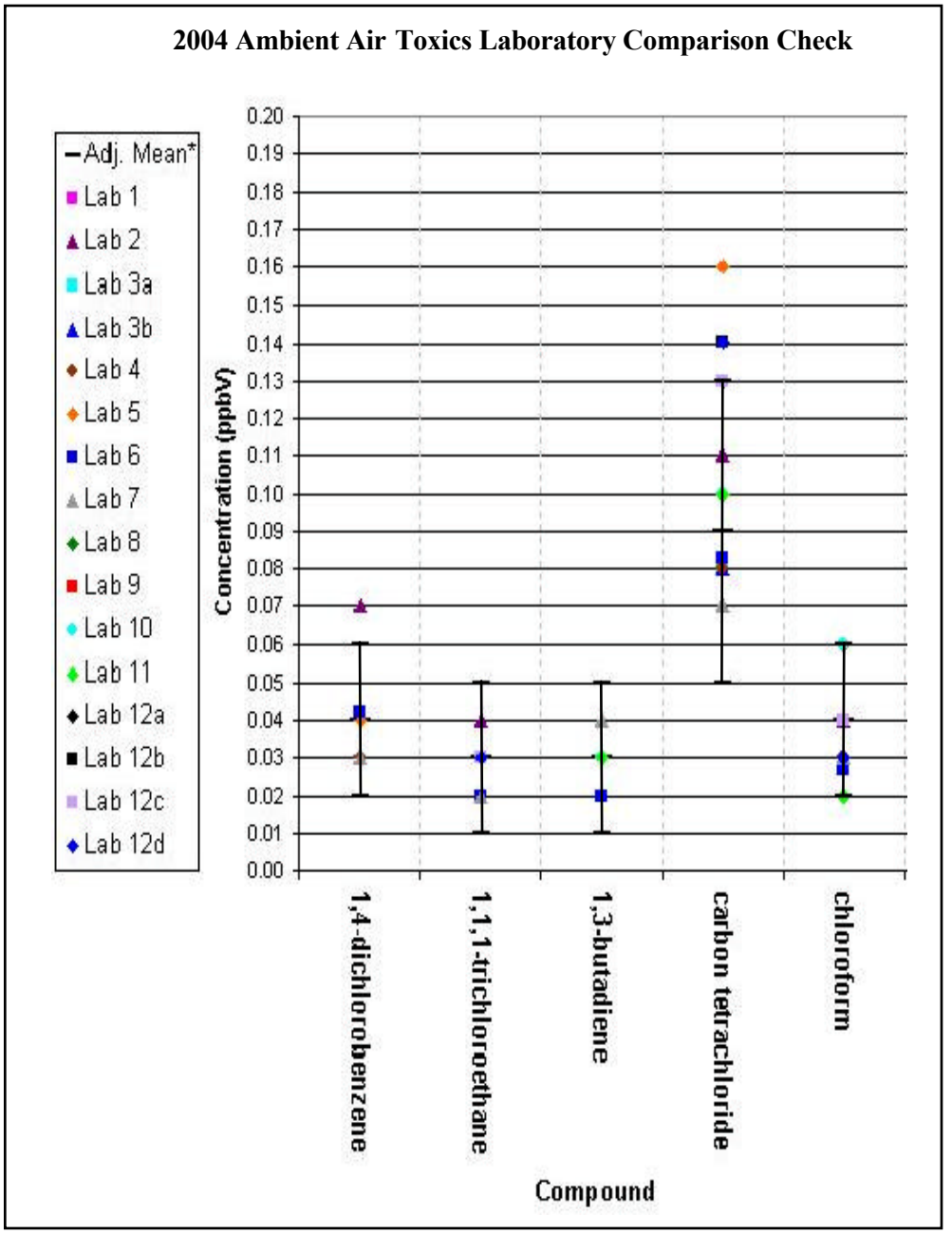


Figure 3.



Developing Criteria for Equivalency Status for Continuous $PM_{2.5}$ Samplers

Basil W. Coutant, Battelle

Abstract

EPA has initiated the development of a proposed methodology and criteria for designating a continuous particulate matter (PM) Class III sampler as being equivalent to the Federal Reference Method (FRM) samplers for sampling $PM_{2.5}$. Because a wide variety of possible candidate sampling methods may be considered, testing procedures and performance requirements must be individually designed and adapted for the specific sampler method of interest. U.S. EPA regulations state that, upon request, the specific test procedures and performance requirements for each Class III candidate method should be determined on a case-by-case basis. However, for consistency, EPA has developed some guidelines for the performance requirements. These guidelines are driven by the data quality objectives for the $PM_{2.5}$ sampling network.

Once achieving equivalence status, the continuous samplers (uniquely specified by manufacturer, brand, and model number) could provide data for multiple uses: daily reporting, forecasting, within-day monitoring of air quality, event monitoring, and attainment demonstration. With the exception of the last item, these uses are generally not feasible with FRM samplers. The costs associated with the operation of the FRM samplers led to requiring only every sixth day sampling in the main network. Since the other uses mentioned are desired from the samplers, the continuous methods are frequently operated in conjunction with the FRM samplers. It is anticipated that the utilization of continuous samplers with equivalence status will result in significant cost savings for State and Local agencies and provide multiple use data.

To determine the equivalency of a candidate continuous sampler relative to the FRM sampler, daily concentration data need to be obtained from $PM_{2.5}$ samples collected from collocated candidate and FRM samplers at multiple sites. This paper describes Battelle's statistical considerations in the development of the guidance criteria from the current DQOs, under U.S. EPA contract number 68-D-02-061. These considerations build on the work begun by P. Mosquin and F. McElroy, under U.S. EPA contract number 68-D-00-206.

Introduction

This paper describes the proposed methodology and criteria for labeling a continuous particulate matter (PM) sampler as being *equivalent* to Federal Reference Method (FRM) samplers for sampling $PM_{2.5}$. The methodology employed adheres to EPA regulatory requirements for designation of federal reference and equivalent methods for $PM_{2.5}$ sampling as defined in 40 CFR Part 53. Under 40 CFR Part 53, continuous samplers are defined as a Class III candidate method of sampling:

> *Class III equivalent methods include any candidate $PM_{2.5}$ methods that cannot qualify as either Class I or Class II. This class includes any filter-based integrated sampling method having other than a 24-hour $PM_{2.5}$ sample collection interval followed by moisture equilibrium and gravimetric mass. More importantly, Class III also includes filter-based continuous or semi-continuous methods, such as beta attenuation instruments, harmonic oscillating element instruments, and other complete in situ monitor types. Non-filter-based methods such as nephelometry or other optical instruments will also fall into the Class III category.*

Because a wide variety of possible candidate sampling methods fall into the Class III designation, testing procedures and performance requirements must be individually designed and adapted for the specific sampler method of interest. In 40 CFR Part 53, U.S. EPA regulations state that the specific test procedures and performance requirements for each Class III candidate method should be determined on a case-by-case basis upon request, in connection with each proposed or anticipated application for a Class III equivalent method determination. This paper provides details and guidance on the proposed criteria for establishing equivalence of a continuous $PM_{2.5}$ sampler with FRM samplers.

Additional related information on the National Ambient Air Quality Standard (NAAQS) and reference method for $PM_{2.5}$ are defined in 40 CFR Part 50 and the network requirements for surveillance of ambient air quality at State and Local Air Monitoring Stations (SLAMS) are detailed in 40 CFR Part 58.

Sampling Requirements

To determine the equivalency of a candidate continuous sampler (uniquely specified by manufacturer, brand, and model number) relative to the FRM sampler, daily concentration data need to be obtained from $PM_{2.5}$ samples collected from collocated candidate and FRM samplers at multiple sites. To support an equivalence evaluation, the data collection process involving collocated samplers needs to adhere to the following requirements:

- Three (3) to five (5) candidate samplers will be collocated with three (3) FRM samplers. This number of FRM samplers is consistent with existing requirements and improves on the ability to identify statistical outliers in daily concentrations.

- Within a given season of the year, each sampler will be run daily for a target of 30 days (with at least one site having samples collected in multiple seasons).

- On a given day, the required sample collection period for each sampler will be a minimum of 22 hours. Although the recommended sampling period for FRM samplers is 23 to 25 hours, EPA has lowered the minimum sampling period to 22 hours for purposes of this equivalency evaluation in order to allow sufficient time to change out samplers and to perform necessary maintenance between sample runs.

- On a given day, valid data must be available for at least two (2) FRM samplers and at least two (2) candidate samplers in order for any data associated with the day's sample collection to be used in the equivalency evaluation.

- Each sampler at a given site will produce valid measurements on at least 75 percent of the sampling days in a given season. For a 30-day sampling period, this corresponds to a minimum of 23 days per season.

- The acceptable concentration range of sample data is 3 $\mu g/m^3$ to 200 $\mu g/m^3$. Although previous $PM_{2.5}$ method designations had a minimum concentration requirement of 10 $\mu g/m^3$ due to concerns about large variations in measurements as one approached zero as well as a desire to test at higher concentrations, recent experience has shown that testing at the higher range is not as representative as it used to be, and fairly repeatable concentrations can be obtained at low concentrations. Thus, the minimum of 3 $\mu g/m^3$ is consistent with proposed edits of 40 CFR Part 58, Appendix A.

Data collection will be replicated at multiple sites to ensure that the sampling is representative of different aerosol types. Furthermore, for at least one site, sampling will occur in at least two distinct seasons of the year. The above sampling requirements will hold across seasons for each site.

Conditions for Equivalency of a Candidate Sampler

From the daily sample concentration data to be collected from the collocated samplers at a given site, the following four essential measures will be calculated:

- Precision
- Correlation
- Multiplicative bias
- Additive bias.

A candidate sampler needs to achieve specified criteria placed on each of these four measures in order to be classified as equivalent to the FRM sampler. Values for these four measures are calculated separately for each site, and the candidate sampler needs to achieve the specified criteria at each site.

The explicit formulas presented below specify how each of these four measures is calculated, and the criteria that the calculated measures need to satisfy are presented with the formulas. Detail is provided in the following section on the derivation of these formulas and the determination of the equivalence criteria. In calculating the four equivalency measures, true daily $PM_{2.5}$ concentrations at a given site are estimated from the daily means associated with the FRM samplers.

Precision: The precision associated with the candidate sampler data is calculated as:

$$Cand\_prec = \sqrt{\frac{\sum_{i-1}^{D}(Cand\_daily\_CV_i)^2}{D}} \quad . \qquad \text{(Eq. 1)}$$

where the summand is the daily coefficient of variation among the candidate samplers. The precision of the candidate sampler data must be **no greater than 15 percent** in order for the candidate sampler to qualify for equivalency classification.

Correlation: Correlation in the daily means between the FRM and candidate samplers is calculated as follows:

$$r = \frac{D\sum_{i=1}^{D}(FRM\_daily\_mean_i \cdot Cand\_daily\_mean_i) - \left(\sum_{i-1}^{D} FRM\_daily\_mean_i\right)\cdot\left(\sum_{i=1}^{D} Cand\_daily\_mean_i\right)}{D^2 \cdot Cand\_RMS \cdot FRM\_RMS} \qquad \text{(Eq. 2)}$$

The value of this correlation **must exceed the following lower bound**, determined by the value of the coefficient of variation of the daily means of the FRM samplers (denoted CCV), in order to qualify for equivalency classification.

$$Correlation\ lower\ bound(CCV) = \begin{cases} 0.93 & if \quad CCV < 0.3 \\ 0.87 + 0.2 \cdot CCV & if \quad 0.3 \le CCV < 0.4 \\ 0.95 & if \quad 0.4 \le CCV \end{cases} \qquad \text{(Eq. 3)}$$

Thus, the value of this lower bound is determined by the value of CCV. It equals 0.93 when CCV is no higher than 0.3, increases linearly from 0.93 to 0.95 as the value of CCV increases from 0.3 to 0.4, then equals 0.95 for all values of CCV above 0.4.

Multiplicative bias: The multiplicative bias is the slope of the ordinary least-squares line between the daily means of the candidate and FRM samplers. It is calculated as the correlation (Equation 2) multiplied by the ratio of the root-mean-square deviations from the overall means of the daily means for the candidate and FRM samplers:

$$b = r \cdot \frac{Cand\_RMS}{FRM\_RMS} \quad . \qquad \text{(Eq. 4)}$$

The multiplicative bias associated with the candidate sampler **must fall between 0.90 and 1.10** in order for the candidate sampler to qualify for equivalency classification.

Additive bias: The additive bias is the intercept of the ordinary least-squares line between the daily means of the candidate and FRM samplers. It is dependent on the daily means associated with the candidate sampler, the overall mean (associated with the FRM samplers), and the calculated value for multiplicative bias (Equation 4). The formula for additive bias is as follows:

$$a = \left(\frac{1}{D} \cdot \sum_{i=1}^{D} Cand\_daily\_mean_i \right) - b \cdot Overall\_mean \qquad \text{(Eq. 5)}$$

The additive bias associated with the candidate sampler **must fall between $a_1(b)$ and $a_2(b)$** in order to qualify for equivalency classification. These lower and upper bounds, which are linear functions of the multiplicative bias (b), are as follows:

$$a_1(b) = 15.05 - 0.92(18.8)b = 15.05 - 17.31b \qquad \text{(Eq. 6)}$$

$$a_2(b) = 15.05 - 1.08(12.2)b = 15.05 - 13.20b \qquad \text{(Eq. 7)}$$

Considerations in the Development of the Criteria

The overall format of the criteria was decided in advance of the project; i.e., it was decided ahead of time that the criteria should be developed based on a linear regression between the daily means of collocated candidate and FRM samplers. The purpose here was to develop the exact criteria that would be used and to develop them based on the existing DQOs for $PM_{2.5}$. This desire to make the criteria tied to the existing DQOs causes a problem, because ordinary regression methods are based on assumptions that are in conflict with the assumptions behind the DQOs. Since the assumptions behind the DQOs are based on observed properties of $PM_{2.5}$ data, the ordinary regression hypotheses are the ones that need to be modified. Chiefly the expected correlation needed to be reevaluated.

Expected Correlation and Candidate Sampler Precision

Correlation is a measure of how well a linear error model fits the data. While unlikely, samplers with seasonal differences in how well they predict the FRM measurements could fail to meet the correlation requirement even though they may have total square errors less than the amounts implied by the other requirements. This is, in fact, a desirable characteristic of the criteria. The reason is that significant seasonal effects would require extensive testing (at least a year per site) to assure that the sampler would always be within a reasonable range of an FRM. On the other hand, consistent biases are much easier to estimate and control. Given that the aim is to have equivalency testing criteria that can be done with much less than a year of data, this property is desirable.

The expected correlation between the daily means of n_1 reference samplers with precision σ_r, n_2 candidate samplers with precision σ_c and over a period with a population CV of τ_u can be written as follows:

$$\frac{1}{\sqrt{\left[1+\sigma_r^2\frac{1}{n_1}\cdot\left(1+\tau_u^{-2}\right)\right]\cdot\left[\left[1+\sigma_c^2\frac{1}{n_2}\cdot\left(1+\tau_u^{-2}\right)\right]\right]}} \qquad \text{(Eq. 8)}$$

This formula is derived in Mosquin and McElroy (2004). Notice that this formula is dependent on the population CV.

Figure 1 shows a plot of the expected correlation for three reference samplers with a precision of 5 percent and three candidate samplers with a precision of 15 percent over a range of population CVs. This shows the sharp dependency on the population CV below 40 percent. Similar plots for different numbers of FRM samplers show that the graph is insensitive to the number of FRM samplers. Hence, the requirement for the number of FRM samplers was set to three rather than five as originally proposed.

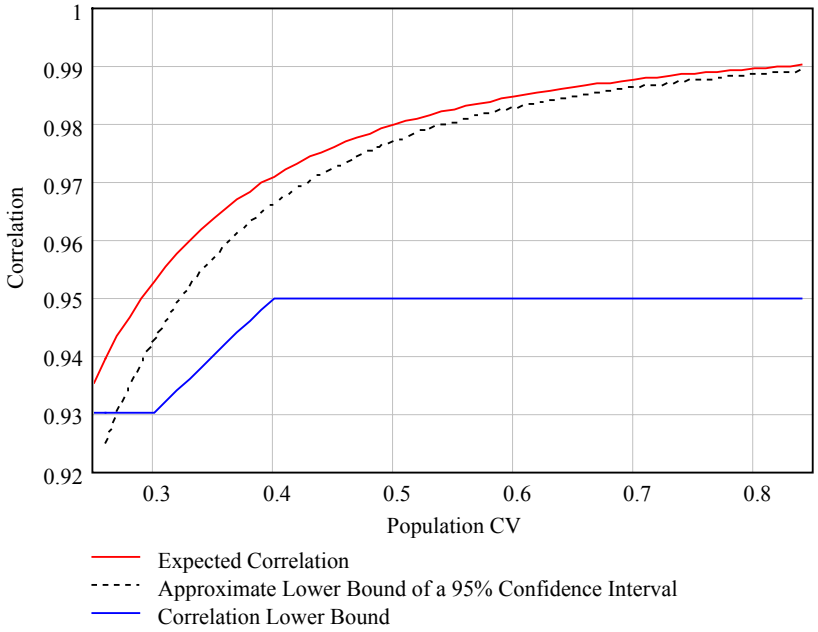


Figure 1. Expected correlation between 46 daily means of three FRM samplers and three candidate samplers with a precision of 15 percent.

Figure 2 shows that for the majority of sites, 30 percent is a reasonable lower bound on the population CV that to be expected. Increasing the measurement imprecision to 20 percent in Figure 1 increases the drop between 30 and 40 percent in the expected correlation. This is one of the reasons for setting the limit on the candidate precision to, at most, 15 percent. The reason for

the limit is not driven by how well the NAAQS decision can be made, but rather how well one can estimate the other parameters.

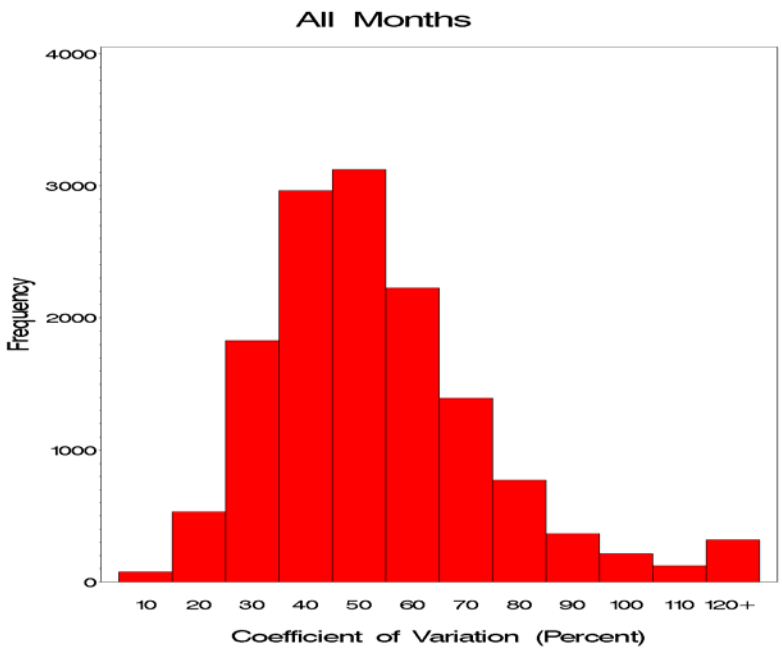


Figure 2. Monthly Coefficients of Variation calculated from data collected in 2003 and downloaded from AQS the first week of July 2004.

Figure 1 also shows the proposed lower limit on the correlation (Equation 3). For a fixed sampler precision it will be both to EPA's and the manufacturer's benefit to choose sites for equivalency testing that are expected to have a large population CV (above 40 percent). As can be observed from Figure 1, this will provide the manufacture additional "acceptable" error in meeting the correlation criterion. Further, as will be seen next, it will improve the bias estimates for EPA.

The approximate lower bound of a 95 percent confidence interval shown in Figure 1 is based on the following transformation of the correlation, r:

$$transform\ (r,n) = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}} \qquad \text{(Eq. 9)}$$

where n is the number of sample pairs. This transformation is approximately distributed as a Student t-distribution with n-2 degrees of freedom. (Note that both the correlation estimate and the above transformation ignore the fact that the FRM samplers have a small amount of measurement error.)

The Acceptance Region for the Additive and Multiplicative Bias

In Mosquin and McElroy (2004), a proposed acceptance region was derived for the additive and multiplicative bias components. The proposed region was based on 1-in-6 day sampling in order to account for increased error rate by local agencies. Because this was felt to be too conservative, a corresponding region for 1-in-3 day sampling was derived and is presented below. This is still conservative since the whole point is to have equivalent measurements collected on a daily or more frequent schedule. The idea behind all of these is that the original DQOs were for 1-in-6 day sampling. By collecting data more frequently, less precise and even more biased data can be used to make decisions that are of the same quality.

The objective was to find combinations of additive and multiplicative bias that would yield the same gray zone that had been established for the 1-in-6 day FRM samplers (i.e., 12.2 $\mu g/m^3$ to 18.8 $\mu g/m^3$), but with a higher sampling rate. First, the *DQO Companion* software tool (Battelle, 2003) was used to derive a gray zone for 1-in-3 day sampling assuming a 10 percent measurement CV and absolute bias. (All parameters were left at the default values except for the sampling rate.) This yielded a gray zone that ranged from 12.64 $\mu g/m^3$ to 18.16 $\mu g/m^3$. (Because the software tool uses random simulations, results will vary among different runs.) From these results, the 5th and 95th percentiles of the simulations were obtained by solving the following two equations for x and y:

$$15.05 - 18.16(0.9)x = 0 \qquad \text{(Eq. 10)}$$

$$15.05 - 12.64(1.1)y = 0 \qquad \text{(Eq. 11)}$$

This yields x = 0.92 and y = 1.08.

Consequently, the same simulations can be used to derive the following acceptable range for the additive bias:

$$a_1(b) = 15.05 - 0.92(18.8)b = 15.05 - 17.31b \qquad \text{(Eq. 12)}$$

$$a_2(b) = 15.05 - 1.08(12.2)b = 15.05 - 13.20b \qquad \text{(Eq. 13)}$$

This resulted in a gray zone of 12.2 $\mu g/m^3$ to 18.8 $\mu g/m^3$ for a given multiplicative bias b (Mosquin and McElroy, 2004)[1].

EPA has set the bounds on the multiplicative bias, b, to be ±10 percent from 100 percent (i.e., $0.9 \leq b \leq 1.1$). Under this range, the box in Figure 3 represents the region of values for multiplicative and additive biases that will correspond to gray zones that range from 12.2 $\mu g/m^3$ to 18.8 $\mu g/m^3$ for 1-in-3 day sampling. The gray zones for the daily sampling with the continuous instrument will be tighter.

[1] The calculations to obtain the functions in Equations 12 and 13 were done with the full precision and rounded at the end, not at the intermediate steps shown.

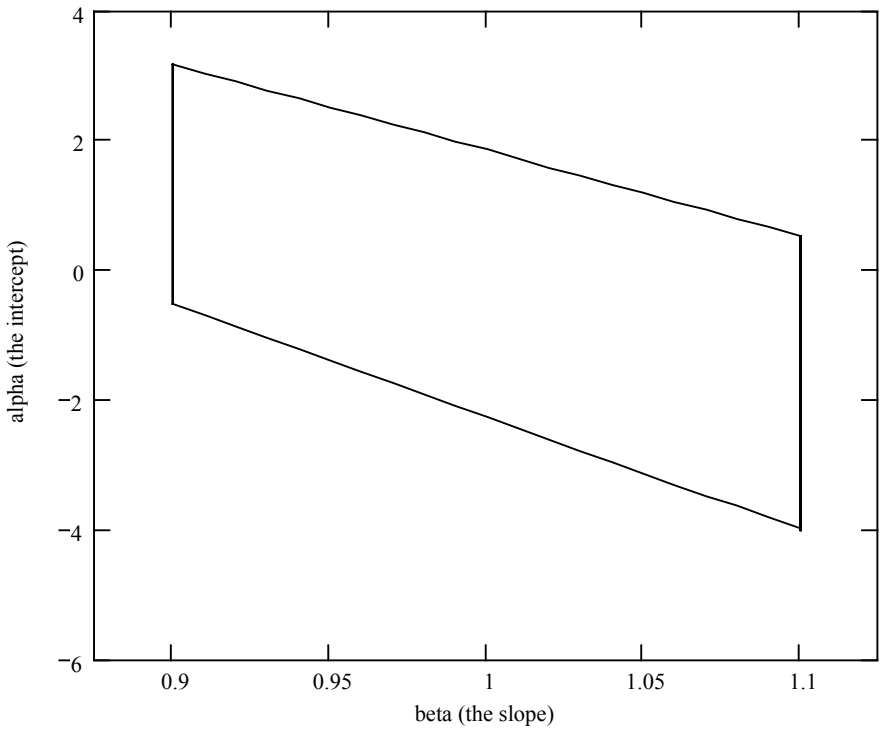


Figure 3. Acceptance range for the additive and multiplicative bias components.

Acknowledgements

The author thanks Tim Hanley and Robert Vanderpool of the U.S. EPA for their valuable input to this project. Also acknowledged is the support of the U.S. EPA under Contract Number 68-D-02-061 for this work, and the work of Paul Mosquin and Frank McElroy, under U.S. EPA Contract Number 68-D-00-206. Mr. Mosquin and Mr. McElroy provided much of the technical foundation for the project.

References

Battelle (2003). "*DQO Companion*, Version 2.0." Software developed and tested for the Environmental Protection Agency Office of Air Quality Planning and Standards under EPA Contracts 68-D-98-030 and 68-D-02-061, September.

Casella, G., and Berger, R.L. (1990). Statistical Inference. Belmont, California: Duxbury Press.

Coutant, B.W., and Holloman, C.H. (2003). "Statistics for Criteria Pollutant Data Quality Objectives." Battelle technical report to U.S. EPA, Contract No. 68-D-02-061, September 15.

Mosquin, P., and McElroy, F. (2004). "Data Quality Objectives for PM Continuous Methods II." ManTech technical report to the U.S. EPA, report number TR-CAN-04-02, June.